Volume 2 *April, 1932* *Number 4*

PHYSICS

THE NEW SPECTROSCOPY LABORATORY OF THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY

By Karl T. Compton

(Received March 1, 1932)

ONE of the features of interest at the recent Cambridge meetings of the American Physical Society and the Optical Society of America was an inspection of the new Spectroscopy Laboratory of M.I.T. This laboratory has been built in conjunction with the erection of the new Physics-Chemistry Reasearch Laboratories and expresses a definite conviction that the science of spectroscopy is destined in the future to play a role of increasing importance in bringing about an understanding of the electrical structure of atoms and molecules and of the interatomic and molecular forces which underlie the interpretation of physical and chemical phenomena. Since the general plan of the main laboratories did not permit the convenient installation of long focus grating mountings, it was decided to build a separate adjoining building, designed particularly for spectroscopy and incorporating every feature which might add to its convenience and effectiveness for such work. It is believed to be the only laboratory of this type in existence.

The accompanying figures give a general idea of the plan of the laboratory. The building is 94 ft. 7 in. long and 54 ft. 1 in. wide outside, and stands about 21 ft. above ground, exclusive of the canopy above the roof. It is constructed in three distinct parts: (1) the outer shell and corridor which are supported upon braced wooden piles and form the housing for the isolated laboratory floors; (2) and (3) the two separate isolated floor systems on which the delicate laboratory apparatus is mounted and which are completely isolated from the housing structure.

Structural Features

All parts of the laboratory are supported by means of piles on a thick stratum of compact gray sand. The outer shell and corridor rest on wooden piles and those under the exterior walls are placed in such a way as to form a cut-off against vibrations reaching the inner or apparatus floors. Since the laboratory is in a court entirely surrounded by other buildings, these buildings themselves serve as an additional cut-off for the traffic vibrations, which

are the chief source of difficulty. The separate isolated floor systems on which the apparatus is mounted are supported on 183 precast reenforced concrete piles, including 10 batter or braced piles placed so as to prevent lateral swaying of the isolated floor structures.

The isolated floors are constructed as follows: Reenforced concrete mats 3 ft. in thickness are built upon the concrete piles. Resting on these mats are special vibration-absorbing structures built up in layers in the following order: three 1 in. layers of fine sand, each contained between layers of three-ply roofing felt; a layer of transite board; a 2 in. layer of reenforced concrete; another group of three layers of sand separated by roofing felt; another layer of transite board; another 2 in. layer of reenforced concrete; a 6 in. layer of special vibration absorbing cork; finally, the 8 in. reenforced con-

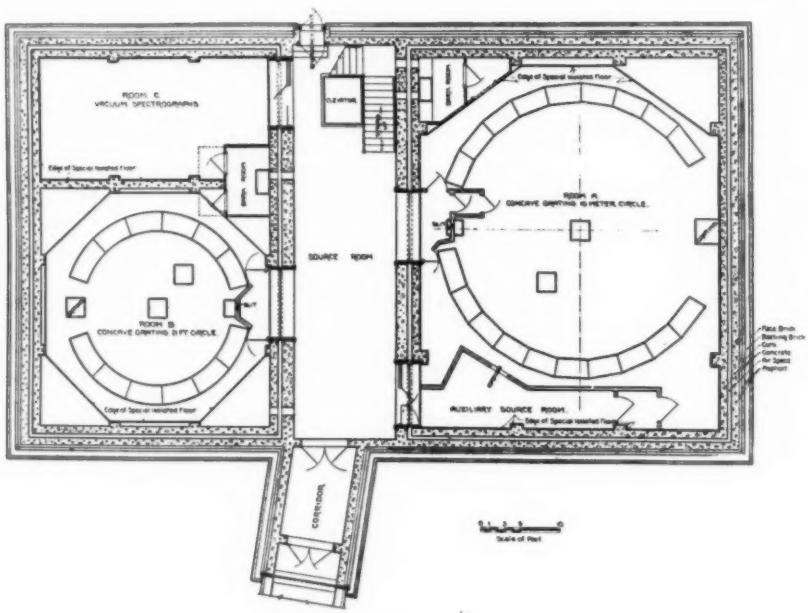

Fig. 1. Spectroscopy Laboratory, first floor plan.

crete floor with smooth cement upper surface. The function of the felt and of the cork is to absorb vertical vibrations, whereas the function of the layers of sand is by their ball bearing action to absorb lateral vibrations. These layers are separated from the adjoining side walls by thick layers of felt. Seismographic tests of the completed structure have shown that lateral vibrations are almost completely absent, and vertical vibrations have been very greatly reduced by this construction.

In order to minimize the effect of outside changes in temperature on the temperature of the rooms, the entire outer wall of the building contains an 8 in. layer of cork embedded between two 12 in. masonry walls. Furthermore, the large grating rooms on the first story, where constancy of temperature is most important, are surrounded by an additional 12 in. concrete

wall with a 6 in. air space. Furthermore, the roof contains 12 in. of cork embedded between thick layers of concrete, and is further protected against changes of outside temperature by having built above it a steel canopy, so constructed as to reflect the sun's rays and to create a circulation of air to diminish the effect of high temperature during the summer. In winter this circulation can be cut off, providing a still blanket of warm air. The entire construction is such that the greatest conceivable change in external temperature could not change the inside temperature at a rate greater than 1° per month.

The vestibule connecting with the main physics research laboratory is provided with three sets of doors, thus providing two air locks. Opening on to the court is an emergency exit which was deemed necessary in view of the fact that there are no other openings in the wall of the building. There is, however, a 6 in. pipe opening in the roof, which is ordinarily closed by a plug but through which noxious gases may be expelled from the building in case experiments are undertaken involving their production.

Laboratory Services

In every research room and at several points in the main corridor (the source room) there are switchboards which make available the following outlets:

240 volt, 3-phase a.c.
240 volt or 120 volt, 1-phase, 60 cycle a.c.
240/120 volt d.c.
700/350 volt d.c.
2 60-ampere lines running from the central distributing switchboard.
8 signal lines running throughout the laboratory.

In addition to these services, the source room is provided with 400-ampere switches for alternating current from a special transformer and for direct current from a special high capacity set of storage batteries. The sources of these currents are all located in the main research laboratory, closely adjacent.

Mounted on vibration-proof bridges, crossing the open space in the second floor between the galleries, are a 200,000 (or 100,000) volt rectified d.c. generator and a 2000 volt motor generator set. High insulation flexible leads from these generators can deliver this power in any part of the laboratory. These generators are mounted on rubber wheels so that they can easily be moved away and the bridges taken up in case it is desired to have full access to the open space between the galleries for the hoisting of large apparatus to the second floor.

Gas, water and ordinary electric service with wall receptacles are provided as usual, care being taken that the pipes and conduits are provided with flexible joints wherever they bridge across from the outer shell of the building to the isolated floors or walls.

Heating and Ventilating

The building is heated and ventilated by air driven in through insulated ducts from an air conditioning plant located in one of the closely adjoining buildings. All air entering the building is conditioned by a standard plant of the type used in theatres, delivering 35 cu. ft. of air per second at 65° ± 2° F and 45 percent humidity. This air enters the source room through seven distributed ducts and leaves through an equal number of ducts, recirculating through the machine with 20 percent of fresh outside air added in each complete cycle. Each room in the laboratory takes its air separately from the source room, having the air blown in by its own fan through a light-proof duct containing a heating element controlled by a thermostat within the room. In this way the source room is kept within 2° of 65° F, and the adjoining rooms can be automatically maintained within ±0.1° F at any desired temperature between about 67° F and 75° F. Automatic temperature registration apparatus permits a complete record of temperature variations to be obtained for each room. A convenient feature in maintaining temperature constancy is the arrangement whereby all electric light switches are so arranged that whenever the lights in a room are turned off an equal load of heat can be turned on.

Laboratory Equipment

In rooms *A* and *B* are mounted the large gratings with modified Paschen mounts arranged so that the complete circles are available. The slits project through floating walls so that the sources of light are shot in from the source room through the slits onto the gratings. Both rooms are also provided with mountings of the Wadsworth stigmatic type which can be used simultaneously.

In room *A* is a 6 in. concave grating of 15,000 lines per inch and 34 ft. radius, while in Room *B* there is a 16 in. concave grating, 15,000 lines per inch, of 21 ft. radius. The focal plane tables consist of albarene stone slabs four inches thick and twenty-eight inches wide supported on and bolted to heavy concrete piers, making, together with the floor, a very rigid structure. Steel sleepers are bolted to these tables about six inches apart, and these sleepers carry the steel rails which hold the cassettes and plate holders. The sleepers are furnished with fine focussing adjustments, and excellent definition and resolving power have been obtained.

Room *C* has been designed for vacuum spectrographs or for grating mountings of the Wadsworth or Meggers type.

On the second floor, (Fig. 2), are a number of smaller rooms. Room *D* is fitted out for cathode sputtering and miscellaneous operations. Room *E* contains the two meter vacuum spectrograph constructed with the aid of a grant from the Carnegie Institution of Washington and arranged to spread the spectrum from 0 to 2000A over a 24 in. Schumann plate. These spectra are being excited by various methods of controlled electron impact in pure and in mixed gases. Room *F* contains a 21 ft. vacuum spectrograph, the

largest of this type in existence, provided with a 30,000 line per inch glass grating and a gas pumping system of extraordinary speed.

Room *G* is at present being used for the assembly of apparatus but is ultimately designed as headquarters for spectroradiometry.

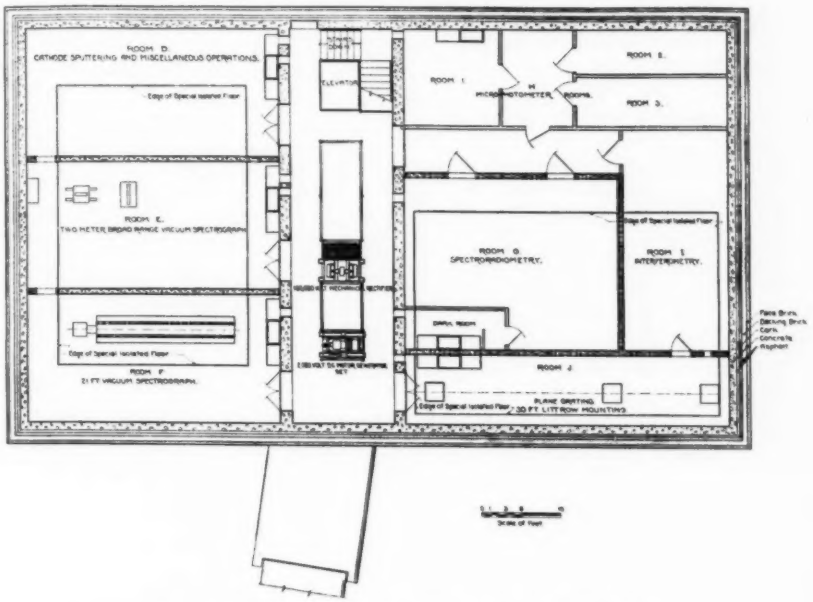

Fig. 2. Spectroscopy Laboratory, second floor plan.

The three rooms in *H* contain three microphotometers for measuring intensities of spectrum lines. The instruments in 2 and 3 are spotting instruments, while that in room 1 is automatically recording.

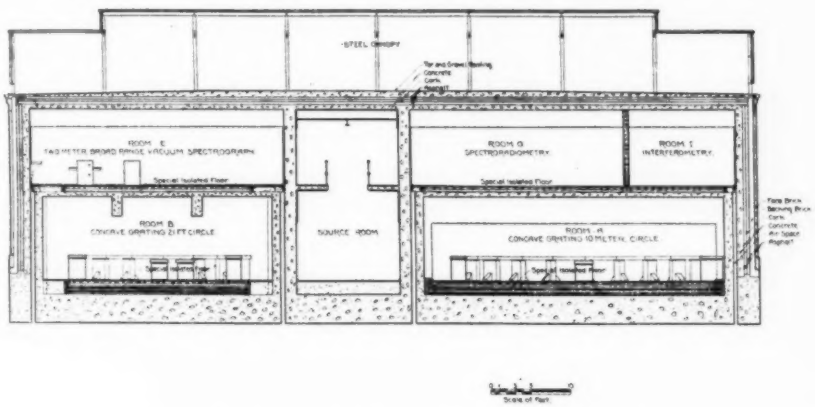

Fig. 3. Spectroscopy Laboratory, longitudinal section.

Room *I* is designed for interferometry but is being used at present for spectroradiometry with monochromators of glass, quartz and fluorite.

In room *J* is mounted a 6 in. plane grating, 15,000 lines per inch, mounted

in the Littrow manner with a 30 ft. lens. This instrument can be used in all orders up to the fourth, and gives a remarkably high resolving power in the visible region. It is being used in connection with etalons, Lummer-Gehrke plates, and other equipment of high resolving power. This grating has been loaned by Professor Norton A. Kent, of Boston University, who is an honorary research associate of the Institute, and who obtained the grating and lens through the aid of a grant from the Rumford Fund.

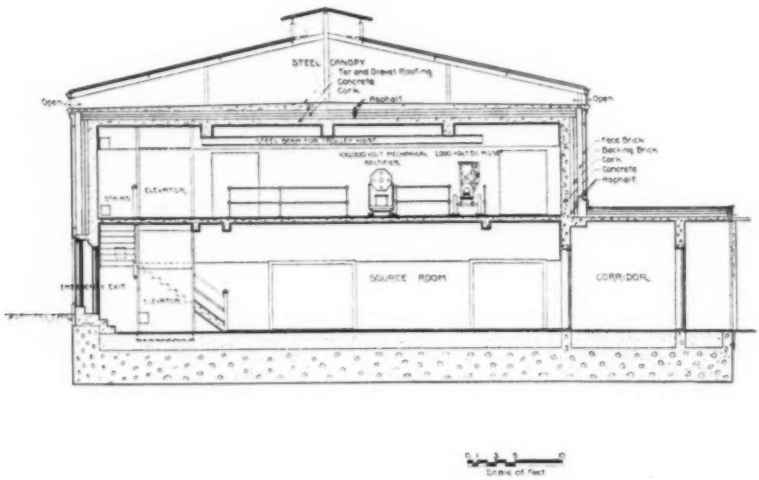

Fig. 4. Spectroscopy Laboratory, transverse section.

In conclusion it may be said that the present work in the Spectroscopy Laboratory emphasizes three main fields: spectral intensity measurements; spectroscopy of the extreme ultraviolet; and spectroscopy at extremely high resolving power. While the equipment now on hand has been set up with the needs of these special fields primarily in mind, the building has been so constructed and the apparatus has been set up in such a form as to allow sufficient flexibility for taking up other important spectroscopic fields when this may seem desirable.

The engineering architects of this building are Chas. T. Main, Inc. of Boston, and the builders are Stone and Webster, Inc., who also built the main educational plant of the Institute.

AN X-RAY SPECTROMETER WITH STATIONARY ION CHAMBER

By Rudolf C. Hergenrother
California Institute of Technology, Pasadena, California

(Received February 2, 1932)

Abstract

A Bragg type ionization x-ray spectrometer which was designed for making quantitative measurements of integrated x-ray intensities reflected from crystals is described. The essential feature of the design is that the ionization chamber is fixed in position, thus allowing a short fixed connection between ion chamber and electrometer (Hoffman). The ion chamber and x-ray tube used with the instrument are described.

Introduction

THE x-ray spectrometer which is described below was used for measuring the integrated intensities of x-ray reflections from Bi single cyrstals. The results of these measurments are described in another article.[1]

The usual design of x-ray spectrometer used for such purposes is the one originated by Bragg. This design utilizes an x-ray tube which is fixed in position and has a moveable crystal holder and a moveable ionization chamber. With a moveable ionization chamber it is necessary either to attach the electrometer to the ionization chamber and move the two as a unit, or to have the electrometer fixed near the axis of rotation of the ion chamber and to have a moveable connection between ion chamber and electrometer (Bragg design).

The writer wished to use a Hoffmann[2] electrometer for measuring the ionization current because of the great stability and high sensitivity of that instrument. Since the electrical connection to the needle of the Hoffmann electrometer should pass through an evacuated tube (if the greatest stability is to be attained), it was considered impractical to use a moveable connecton between ion chamber and electrometer. Accordingly, a spectrometer was designed, having the ion chamber fixed in position and having the x-ray tube and crystal holder moveable. The use of a short stationary connection between ion chamber and electrometer instead of a long moveable connection reduces the electrical capacity to a minimum and eliminates the possibility of a variation of the capacity of the connection itself.

The Spectrometer

Fig. 1. shows a drawing which illustrates the essential features of the spectrometer. The central post A is fixed to the base B. The upper part of this post, which is tapered on the outside, carries a circular angle scale C and a moveable arm D. The scale C, which is marked in intervals of one degree at

[1] A Goetz and R. C. Hergenrother, Phys. Rev. **39**, (in press) (1932).
[2] G. Hoffmann, Ann. d. Physik **42**, 1196 (1913); **52**, 665 (1917).

the edge of its upper surface, is set to an arbitrary fixed position on the post *A*. The moveable arm *D* carries on one end the x-ray tube *E* with the slits *F* and *G*, and on the other end a counter weight *R*. An indicator having a ver-

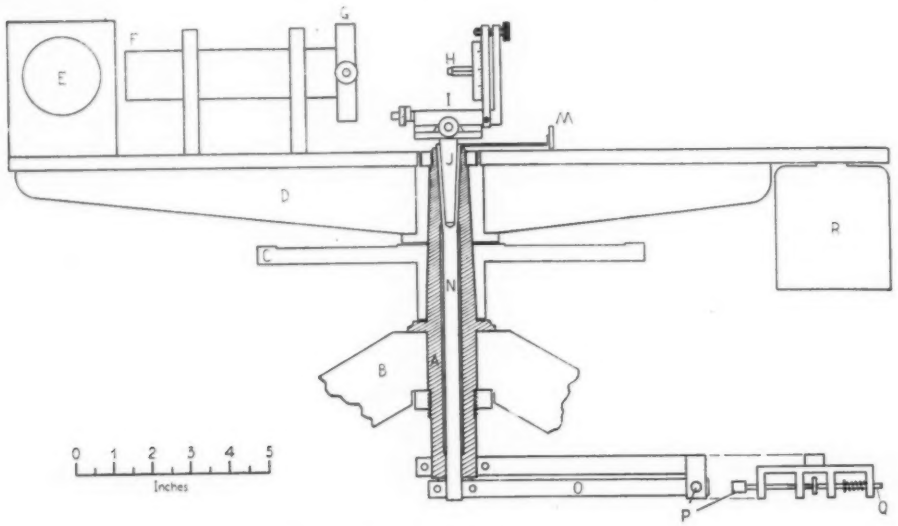

Fig. 1. Drawing of Spectrometer.

nier scale is also fastened to arm *D* in order to measure its position relative to to the fixed scale *C*. This is shown in detail in Fig. 2.

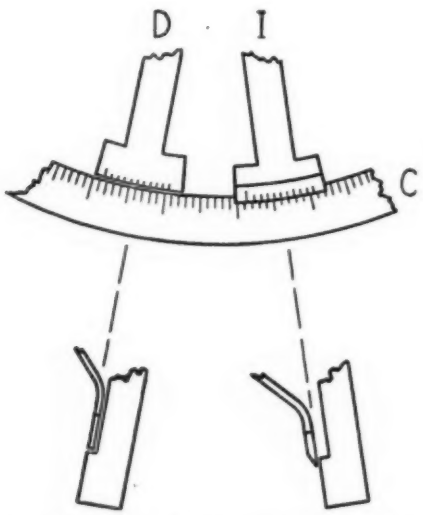

Fig. 2. Detail of vernier indicators. *D* measures the position of the x-ray tube, and *I* measures the position of the crystal. The lower figures show respective cross sections of the indicators.

The crystal *H* is mounted on the crystal table or goniometer *I* which is shown in detail in Fig. 3. The base of this goniometer is fastened to a cone *J*

which fits into the conical hole in the upper part of the central shaft N. This allows the goniometer to be removed from the spectrometer as a unit and also allows the goniometer to be set at any arbitrary position relative to the shaft N. The base of the goniometer also carries a vernier indicator which

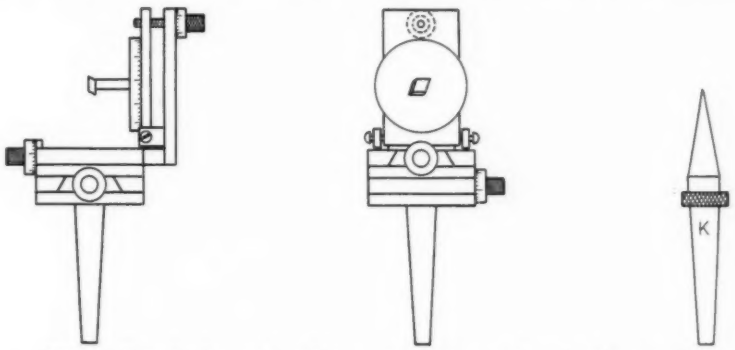

Fig. 3. Detail of crystal table or goniometer and centering cone K.

measures its position relative to the fixed scale C. Fig. 2 shows how the two indicators already mentioned are arranged so that they use the same angle scale without interferring with each other.

The central shaft N (Fig. 1) whose upper end fits into a conical hole in the upper end of post A, has the arm O attached to its lower end. The end of arm O is held between the micrometer screw P and the spring operated plunger Q.

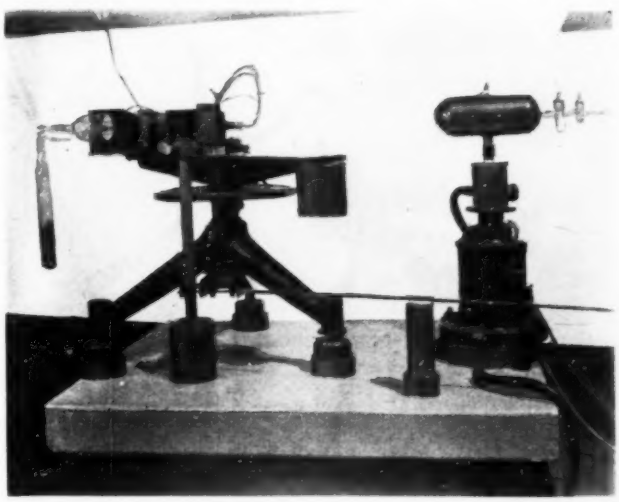

Fig. 4a. Photographs of the spectrometer. The lead shields have been removed from the ion chamber.

This allows a sensitive adjustment of the angle of the goniometer over a range of 4 degrees. The angular displacement of the upper end of N from its mean position is measured by means of mirror M which reflects the light from a fixed scale to a telescope. It is necessary to attach mirror M to the upper end

of shaft N in order to make the reading of the angular position of the goniometer independent of any twist in shaft N. Fig. (4) shows a photograph of the spectrometer.

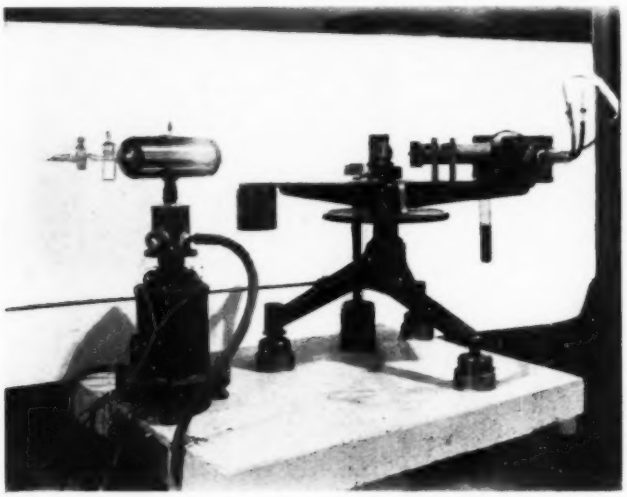

Fig. 4b.

The X-ray Tube

The x-ray tube should preferably be of such design that the central part of the tube may be brought within 2 or 3 inches of the supporting arm D. A Meuller Metalix tube (which has the central part made of metal) was originally used. Subsequently a special tube which is illustrated in Fig. 5 was built and successfully used with the instrument.

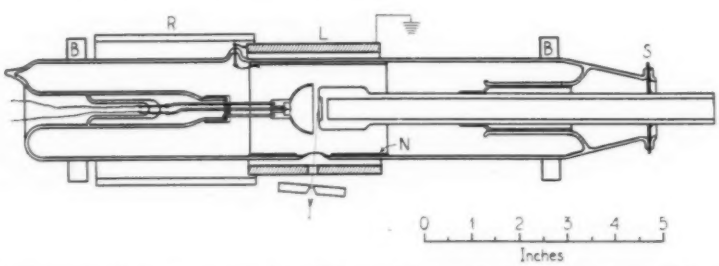

Fig. 5. Cross section of x-ray tube. N is a nickel cylinder which fits inside of the glass wall. The glass wall is made thin, as indicated, in the region where the x-rays which go through the slit, pass through the glass. The nickel cylinder has an opening in the corresponding position. The lead sheath L is electrically connected to N and grounded. R is a sheath of x-ray opaque (lead salts impregnated) rubber. B-B are bakelite supports. S is a copper disk seal.

The Ion Chamber

The ionization chamber which is shown in Fig. 6 consists of a Pyrex glass cylinder with rounded ends, the inside of which is gold plated to form one

electrode. A tungsten wire *A* sealed through the glass, having a small platinum strip welded to it and sealed to the inner surface of the glass, serves to make an outside electrical connection ot the inner gold coating. At one end of the ion chamber is a thin 0.002″) Pyrex window *B*.[3] The opposite end of the cylinder has a tube *C* sealed to it which serves to connect the ion chamber with the P_2O_5 cell and the filling stopcock. The tube *C* is to one side of the center and in line with the collector wire *D*. In assembling the ion chamber, the collector wire *D* is introduced through *C* before the cell and stopcock are sealed on and, since the side of the collector wire toward *C* is in effect a long screw, this is screwed tight against the support *E*, holding the collector wire rigidly in plate. A quartz tube *F* serves as an insulating support for the col-

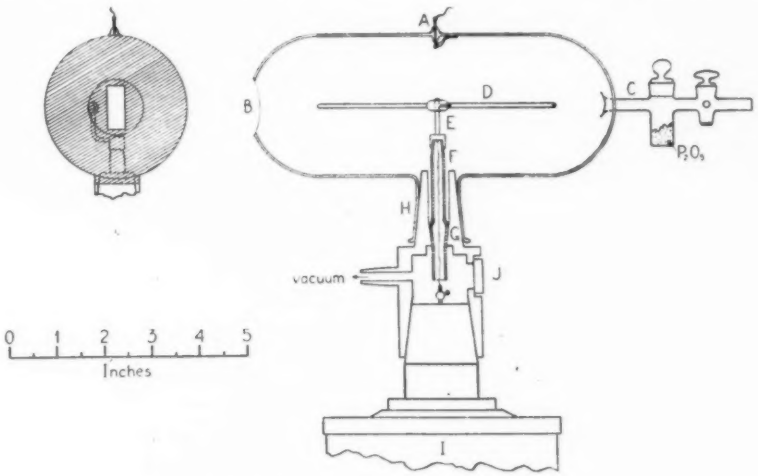

Fig. 6. Detail of ionization chamber.

lector and is fitted by a tapered waxed joint to the grounded metal piece *G* which serves the double purpose of acting as a guard ring and shielding a large portion of the quartz from electric field gradients. This metal piece *G* is also fitted with a tapered waxed joint to the conical opening *H* in the lower central part of the cylinder. The whole assembly is fitted directly on top of the Hoffmann electrometer *I*. A waxed-on window *J* allows access to connection screw of the electrometer.

All metal surfaces which are on the inside of the ion chamber are gold plated to prevent chemical reaction with the methyl bromide with which the chamber is filled.[4]

[3] This was made by opening a round hole at one end of the glass cylinder and sealing a thin Pyrex bulb to the edge. The excess glass was trimmed off and the edge was remelted in the flame and annealed.

[4] The gold coating on the inside of the glass is produced by evaporation in a vacuum of a bead of gold placed in a heated helical tungsten filament. To do this operation, *F* is replaced by a glass tube having two wires sealed through it, to which are attached the ends of the tungsten filament holding the gold bead. This is waxed to *G* which is in turn waxed to *H* after which the chamber is exhausted through *C*. The chamber is mounted vertically with the win-

The ion chamber is surrounded by a heavy lead shield to protect against scattered x-rays. A brass tube with wide lead slits at each end leads from a point several inches from the crystal to the opening in the lead shield in front of the ion chamber window. The purpose of this arrangment is to limit the solid angle which the window subtends outside of the shielding system and thus reduce to a minimum the scattered rays which can enter the window.

Adjustment of Slits

The slits *F* and *G* (Fig. 1) are of the conventional type of x-ray slit having adjustable lead jaws. The slits are lined up with the axis of rotation of the goniometer by means of the centering cone *K* (Fig. 3) which can be inserted in the conical hole in *N* in place of the goniometer. The upper part of *K* is tapered to a point which is accurately machined so that the point lies in the axis of rotation. When *K* is in position, a light is placed in front of *F* and allowed to shine through slit *G* and cast a shadow of cone *J* on a screen placed behind the latter. The slit *G* is then adjusted until the diffraction pattern caused by the point of the cone *J* is centered on the diffraction pattern caused by slit *G*.

Adjustment of Crystal

The crystal is adjusted with its surface in the axis of rotation in the following way. The centering cone *K* (Fig. 3) is set in place and is viewed through a telescope which is adjusted so that a vertical crosshair passes through the point of *K*. The centering cone is then replaced by the crystal goniometer which is adjusted so that the crosshair of the telescope passes through the center of the crystal face when the latter is placed perpendicular to the line of sight and also when it is placed almost tangential to this, so that a small projected area can be seen. The plane of the crystal is adjusted to the vertical by observing in the telescope the reflection from the surface of the crystal of a small light source shining through the center of slit *F* (Fig. 1).

Conclusion

The writer wishes to express his indebtedness to Dr. J. W. M. DuMond for suggestions regarding the design of the ion chamber and to Mr. Julius Pearson who constructed the spectrometer.

---

dow *B* at the bottom so that a thin piece of iron can be placed inside to shield the center of the window from the heavy gold deposit. When the gold has been evaporated sufficiently so that the coating is opaque, the piece of iron is displaced to one side with a magnet, and an additional light coating of gold sufficient to make it electrically conducting is put on the clear area. The advantage of this thin coating on the window is that in addition to electrical shielding, it causes small x-ray absorption and allows one to look into the ion chamber during assembly.

A CATHODE-RAY FREQUENCY MULTIPLIER

By Noel C. Jamison
Northwestern University

(Received January 21, 1932)

Abstract

A vacuum tube was constructed having a filament and slit arrangement for the production of a sheet of electrons, one pair of deflecting plates, and a receiving plate composed of a series of rectangular conductors. Adjacent conductors were insulated from one another and alternate conductors were in electrical connection. An alternating potential difference applied to the deflecting plates causes the electron beam to move across the surface of the receiving plate, current flowing first to one set of conductors and then to the other set. During one complete cycle of the deflecting plate field, the currents from the two sets of conductors will pulsate with frequencies which are multiples of the frequency applied to the deflecting plates. By applying steady potentials to the deflecting plates, the current to each set of conductors may be measured and plotted as a function of the deflecting plate potential difference. Such static characteristic curves were examined for widely different currents, potentials, and gas content of the tube. Tests with an alternating potential difference applied to the deflecting plates gave very definite indications of receiving plate currents of two, four, and six times the frequency of the applied field.

Introduction

At the present time there seems to be no method available for the production of intense, sustained, electric oscillations in the frequency range above 2×10^9 cycles/sec. Okabe[1] has reported comparatively intense oscillations of wave-length 14.5 cm, using a magnetron tube, and very weak oscillations of wave-lengths as short as 3.16 cm. A device which would furnish considerable power and which would permit critical frequency adjustment in the range beyond that to which the ordinary vacuum tube oscillator is adapted would be a useful tool. The frequency multiplier described in this paper was conceived with this object in view.

Such a device, capable of multiplying very high frequencies, would also be useful in obtaining frequency control for vacuum tube oscillators working in the range beyond the region where crystal control may be used.

The object of the present work has been to design a frequency multiplier and to verify experimentally the fundamental principle involved.

Theory and Apparatus

The fundamental principle involved may be used as the basis for design of two slightly different frequency multipliers. While the first type is not so desirable experimentally from some standpoints, it is perhaps simpler in theory.

In this simpler form the construction is, to a large extent, much the same

[1] Okabe, I.R.E. Proc. **18**, 1748 (1930).

as that of a cathode-ray oscillograph. A means for the production of an electron stream of circular cross-section is provided, and two sets of deflecting plates are placed at right angles to each other and symmetrically with respect to the path of the electron stream when undeviated.

In place of the fluorescent screen of the cathode-ray oscillograph is substituted a plane electrode receiving plate in the form of an annulus. The annulus is composed of n sectors, n being an even number; all adjacent sectors are insulated from one another and alternate sectors are connected by conductors. A lead is brought out from each set of sectors and both sets are held at the same positive potential with respect to the filament.

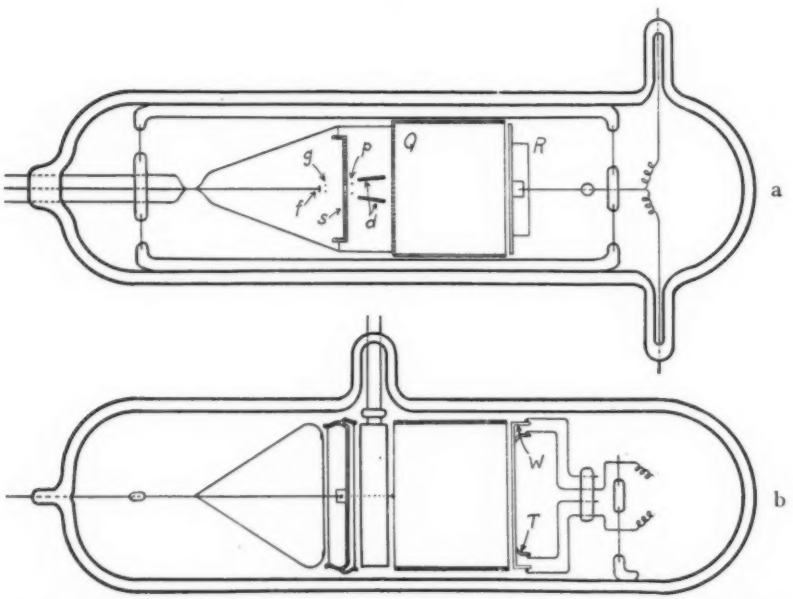

Fig. 1. (a) Cross-section of the tube looking down upon it, showing the filament, f; grid accelerating wires, g; slit, s; field wires, p; deflecting plates, d; shield, Q; and the receiving plate, R, which is composed of two parts, T and W, shown on (b) which is a cross-section at right angles.

Two alternating sinusoidal electric fields of equal intensity and ninety degrees out of phase are set up at right angles by means of varying potentials applied to the deflecting plates. When such a condition obtains, the charge carried by the beam will flow first to one set of sectors and then to the other. Since alternate sectors are electrically connected, during one complete revolution of the beam a pulsating current of $n/2$ times the frequency applied to the deflecting plates will flow in the leads from the annulus. Although the wave-form of the resultant current will only approximate a sine curve, it may be shown that a large portion of the energy would be associated with the fundamental, which would have a frequency $n/2$ times that applied to the deflecting plates.

In order to obtain a large multiplication factor, it would be necessary to

have a large number of sectors and hence an electron beam of small cross-section. The resulting currents would be small.

It is easily shown that should the diameter of the beam be several times the sector width, a pulsating current would still flow in the leads. This mode of operation, however, would not be efficient; power would be sacrificed for the sake of increased multiplication.

The second form of the apparatus as shown in Fig. 1 (a) and Fig. 1 (b) is actually a one-dimensional form of the apparatus already described, using one set of deflecting plates and a sheet of electrons. Phase adjustment in the deflecting plate circuit is thus avoided and much larger currents may be obtained. The annulus is replaced by a series of rectangular conductors insulated from one another, alternate conductors being electrically connected. All experimental results given here were obtained with this type of tube.

Fig. 1(a) shows a cross-section of the tube looking down upon it, and Fig. 1(b), a cross-section at right angles. The oxide-coated filament, f, is

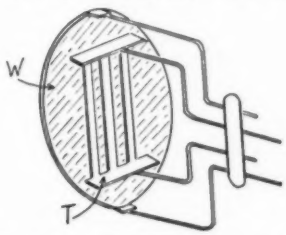

Fig. 2. Detail of receiving plate.

mounted in such a manner that it remains parallel to the grid wires, g, when heated. The grid wires, which furnish the accelerating field, are held taut and in position by hooks welded to the slit plate, s. The wires, p, are supported in the same manner as are the grid wires and are intended to shield the slit region to some extent from the varying potentials applied to the deflecting plates, d. Q is a cylindrical shield with the end near the deflecting plates closed, except for an opening opposite the deflecting plates. The receiving plate, R, consists of two parts, T and W, which are insulated from each other; the construction of the receiving plate is shown in detail in Fig. 2. The parts g, s, p and Q are held at the same potential by internal connections.

The currents and potentials will be designated as follows:

I_f—filament heating current
E_g—potential of the grid and parts of equal potential
I_g—current to the grid and parts of equal potential
E_x—potential difference between deflecting plates
E_d—constant mean potential of the deflecting plates
I_d—current to the deflecting plates
E_r—potential of the receiving plate
I_1—current to part T of the receiving plate
I_2—current to part W of receiving plate.

(All potentials are measured with respect to filament potential.)

With this one-dimensional form of the apparatus, it is possible to obtain very much larger currents than with the form of apparatus first described, with the same length electron path, the same deflecting fields, and the same receiving conductor widths. Since during one complete cycle of the deflecting plate field the beam passes over each conductor in the one-dinensional device twice, the maximum multiplication factor is $(n-1)$, where n is the number of conductors. If the deflecting field is sinusoidal the conductors near the center of the receiving plate must be wider, if the resulting waveform is to have relatively small harmonics. The width of the narrowest conductor is limited by the thickness of the electron beam.

Experimental Results

The most important static characteristic curve of such a tube is that obtained by plotting the currents, I_1 and I_2, as a function of E_x. These curves for the tube described were examined with widely different values of E_g, E_r,

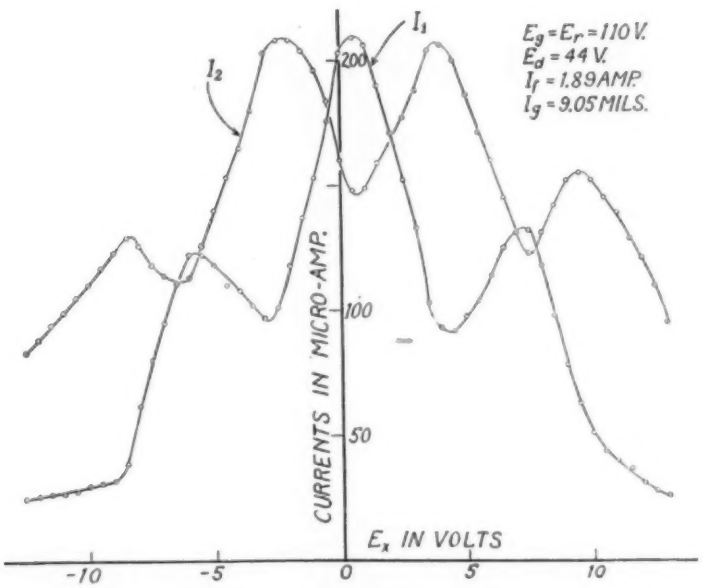

Fig. 3. Static characteristic of the tube in a high vacuum state.

E_d, and I_g and also with different gas content. In case ionization occurs and E_r is less than E_g, both electron and positive ion current will flow to the receiving plate. In an attempt to investigate experimentally the discharge of ions composing the sheath, the conductors constituting the receiving plate were made of equal width.

Fig. 3 shows the receiving plate currents plotted as a function of the deflecting plate potential difference, the tube being well exhausted and carefully chosen potentials being applied to the tube. The data were taken after pressure equilibrium had been established; after taking data for a curve any

reading could be checked within one microampere. As is to be expected, the current, I_1, has three maxima to which correspond three minimum values of I_2. The ultimate decrease of I_2 with increase in potential difference is due to the fact that the deflecting plates and the end of the cylindrical shield, Q, took current when the deflection of the beam was large.

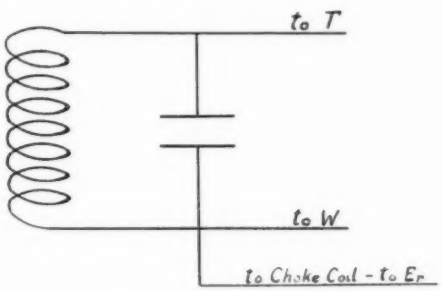

Fig. 4. External circuit.

In a tube with an even number of conductors, and with no loss of current to the deflecting plates for large deflections, the rate of pulsation of current would be the same in each lead to the receiving plate, if a.c. were applied to the deflecting plates. An external circuit of the type shown in Fig. 4 might be used to advantage; the difference $(I_1 - I_2)$ would then assume importance.

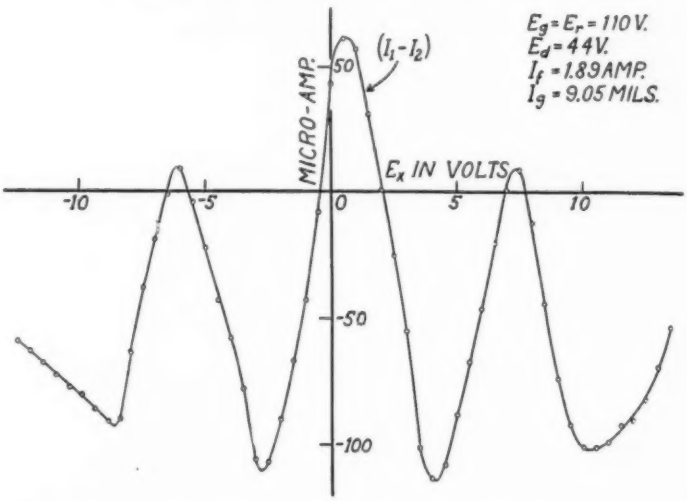

Fig. 5. Current difference, $(I - I_2)$, plotted as a function of the deflecting plate potential difference, as derived from Fig. 3.

Fig. 5 shows this difference plotted as a function of E_x, as derived from Fig. 3, and may be used as the basis for an estimate or calculation of the efficiency of the tube as a frequency multiplier. From Fig. 5 it is evident that the wave-form would be improved considerably by causing the middle of the electron beam to fall on the middle of the central conductor of the receiving plate

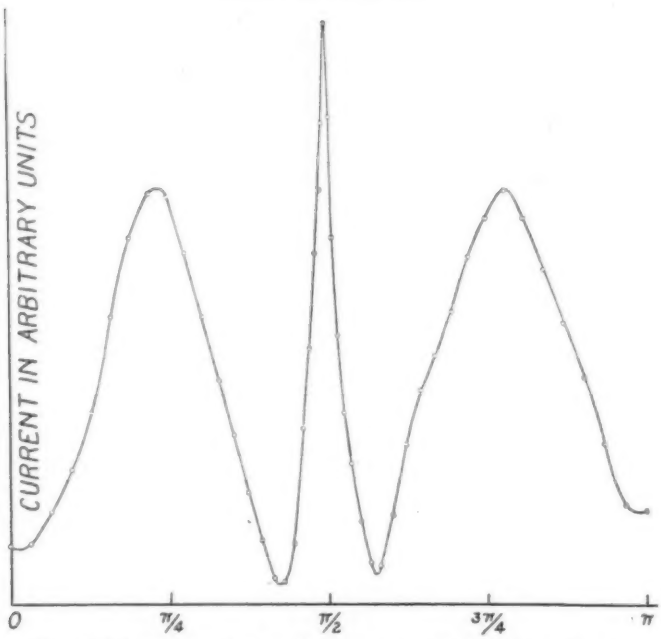

Fig. 6. Wave-form for one-half cycle as derived from Fig. 5.

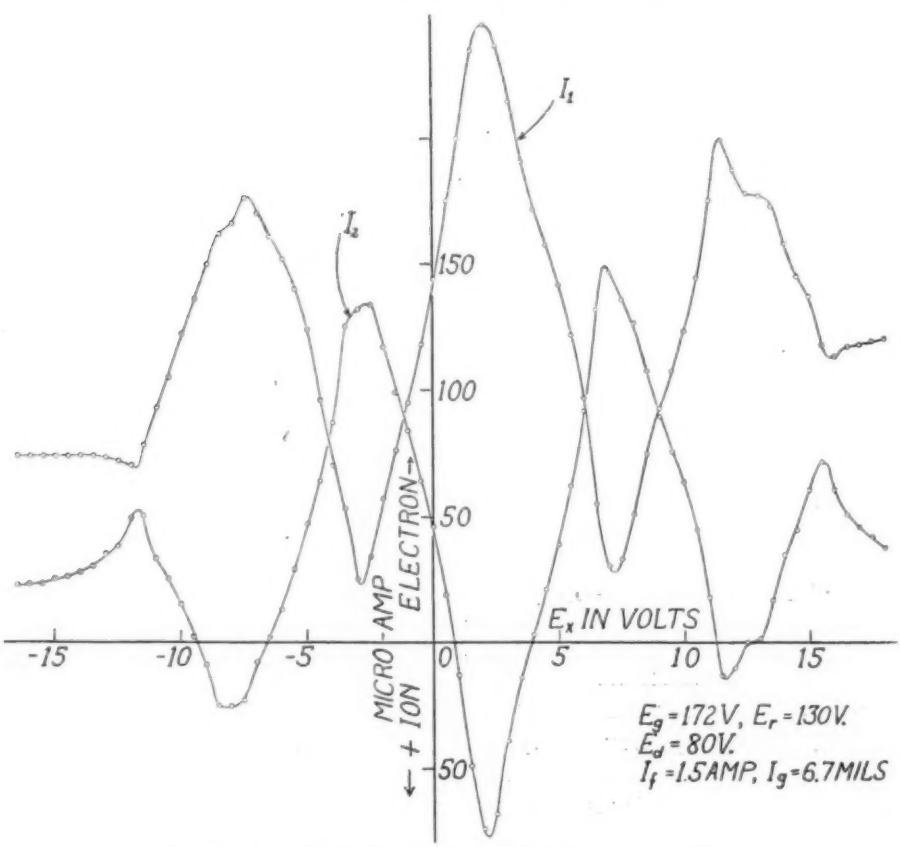

Fig. 7. Static characteristic of the tube in a gaseous state.

with no applied alternating potential difference. This might be accomplished by applying a constant potential difference of about 0.7 volt between the deflecting plates. It may be noted that no adjustment of potentials can be made to cause the beam to pass midway between the deflecting plates if the elements of the tube are not in line. Arbitrarily, the maximum value of the alternating potential difference applied to the tube might be taken as 9.6 volts. Such a choice would favor the production of the sixth harmonic of the frequency applied to the deflecting plates.

Fig. 5 does not show the wave-form of the current that would be supplied to the circuit of Fig. 4; for the electron beam sweeps across the plate with varying velocity. If the alternating potential difference applied to the deflecting plates is sinusoidal, the wave-form of the current may be found by

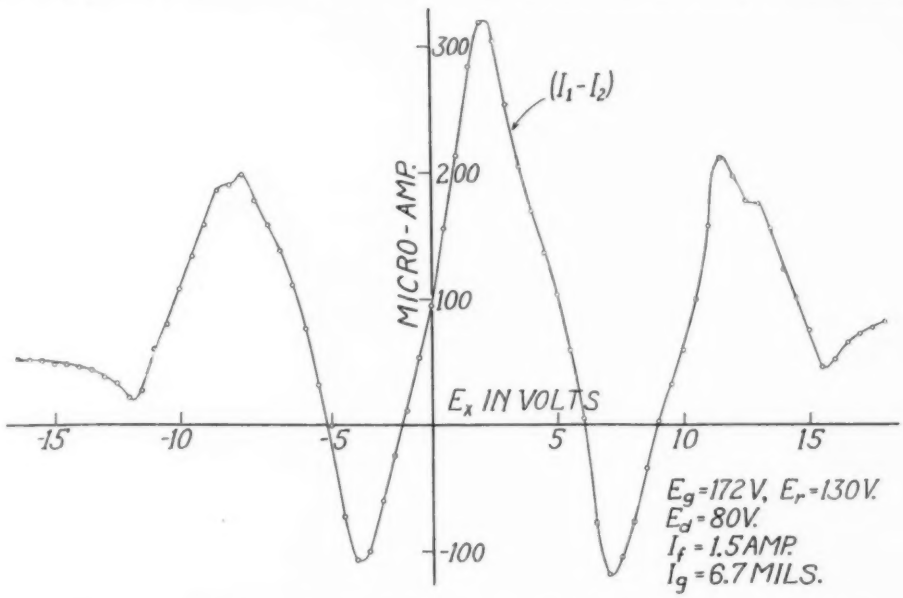

Fig. 8. Current difference $(I-I_2)$ plotted as a function of E_x as derived from Fig. 7.

plotting the difference (I_1-I_2) as a function of E_x, using graph paper that is sinusoidal in one dimension. Fig. 6 shows the wave-form for one half cycle.

Judging from the positions and shapes of the peaks in Fig. 6, one would expect the fourth and sixth harmonics to be strong. If the widths of the conductors were correctly proportioned, a large percent of the available power might be concentrated in one certain harmonic.

Using a circuit such as that shown in Fig. 4, the voltage across the coil was measured roughly with a vacuum tube voltmeter. With the resonant circuit tuned to 2580 k.c. the applied deflecting plate frequency was varied from 400 k.c. to 1300 k.c. Indications of oscillatory current in the resonant circuit were obtained for frequencies 1270, 865, 645, and 432 k.c. of which the frequency 2580 k.c. is the second, third, fourth, and sixth harmonic. The vacuum tube voltmeter showed the second, fourth, and sixth harmonics

to be relatively strong, the third weak, and the fifth was not to be found. In each case the filament current of the multiplier tube was reduced to zero to determine the direct pick-up from the oscillator.

Fig. 7 shows the static characteristic of a tube containing sufficient gas to produce considerable ionization, the receiving plate being held at less than grid potential. In this case both electron and positive ion current flow to the receiving plate. By careful adjustment of the width of the beam, the electron stream may encounter one conductor while positive ions flow to the conductors on either side. Fig. 8 shows the plot of (I_1-I_2) as a function of E_x. Due to the fact that for efficient operation the conductors must be of different widths if the variation of potential of the deflecting plates be sinusoidal, the possibility of utilizing the positive ion current to advantage is doubtful.

A multiplier tube of improved design is being constructed at the present time with which the writer hopes to produce intense oscillations in the frequency range above 2×10^9 cycles/sec.

The writer wishes to thank Professor J. H. Morecroft for advice and criticism and Professor B. J. Spence for his cooperation and continued interest in the problem.

THE CALCULATION OF DETECTION PERFORMANCE FOR LARGE SIGNALS

By J. P. Woods[1]
University of Texas

(Received February 16, 1932)

Abstract

It is mathematically proven that a detector tube which has a broken straight line for its plate current-grid voltage static characteristic will detect the standard radio signal without distortion, regardless of signal strength. The plate detection performance of a tube having a given curve for its static characteristic can be calculated from the extended power series which represents the characteristic. The power series analysis of the curve is obtained from a Fourier series analysis by a method which saves labor and is more accurate than the method of simultaneous equations. Thus, by using twelve or fifteen terms of the power series instead of only the first two or three, and by setting up special formulae, detection performance can be calculated for any signal strength and any modulation. The method is applicable to the tube alone, to a circuit containing the tube and external resistances, and may be applied to any similar non-linear receiving system.

I. Introduction

Difficulty of making calculations for radio circuits

THE calculation of the performance of a radio receiving set, or of any portion of it, is immensely complicated because of the large number of variables and because of the lack of known laws to define the characteristics of some of the most important devices included in the set. As an example of the latter, it has been found necessary to express the relations among the grid and plate voltages and currents of a vacuum tube by means of empirical equations in the forms of power series; there are no satisfactory equations based upon theoretical considerations. Carson first employed such empirical equations in setting up amplification and detection theory. Later, Lewellyn and Chaffee used them.[2]

These equations are usually limited to the square term for two reasons: first, the use of the higher powers makes the calculations very complex; second, for small signals an empirical equation ending with the square term represents the tube characteristic accurately enough. But when plate detection is used, the grid swing covers a large portion of the static characteristic, and the higher powers of the empirical equation must be employed. The result is that the calculation of the performance of a plate detection circuit becomes very complicated.

There is given here a method of calculating plate detection performance under certain limiting conditions. Employing the Carson series method, formulas and tables are developed from which can be calculated the harmonics

[1] Bureau of Engineering Research, The University of Texas, Austin.
[2] See the bibliography.

of the audio-frequency plate current when the standard signal is impressed on the grid circuit. The fundamental output and the amount of distortion are calculated for a detection circuit employing a UX-201A vacuum tube, and the results obtained are compared qualitatively with experimental results obtained by Ballantine.

The standard signal

The radio signal emitted by a broadcasting station consists of a wave of high and constant frequency upon which are superimposed the comparatively low and varying frequencies of voice or music. The audible frequencies are thus transmitted as periodic variations in the amplitude of the high or radio frequency wave. If both high and low frequency waves are sinusoidal, the equation for the signal is

$$e_s = E(1 + m \sin \omega_L t) \sin \omega_H t. \tag{1}$$

The symbols have the significance indicated in the table of symbols at the end of the paper. Actually, the waves of radio transmission and reception are far more complex than that represented by the above expression. However, Eq. (1) lends itself to mathematical manipulation. For that reason, in analyzing the performance of any piece of radio apparatus, it is convenient to assume that the input to the apparatus is represented by this equation. From results obtained using this simplified signal as a standard, it is possible to form an idea of what the performance will be when the input is more complex.

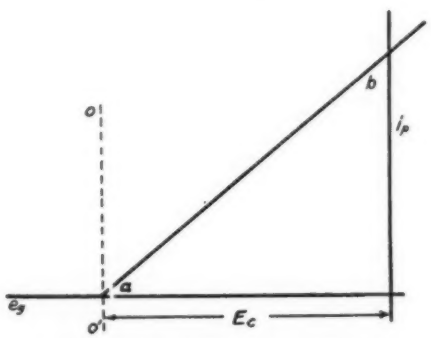

Fig. 1. A linear static characteristic.

II. Ideal Detection

The linear form of the static characteristic

It is generally stated that a vacuum tube circuit having a static characteristic of the form shown in Fig. 1 will give ideal detection when the tube is polarized to the break in the curve. Suppose the input is the standard signal, a modulated wave as drawn in Fig. 2. The static characteristic of Fig. 1 gives linear rectification, and the plate current due to the input voltage e_s is of the form shown in Fig. 3. To demonstrate that the tube is an ideal detector, it is necessary to find the Fourier series expression for the plate current wave of

Fig. 3, and to show that this series contains terms of frequency f_L but none of frequency $2f_L$, $3f_L$, etc.

Derivation of output equation

If, using 0–0′ as an axis, the equation of portion a–b of the static characteristic shown in Fig. 1 is

$$i_p = Ke_s, \tag{2}$$

then the plate current is represented by the positive values of the following equation:

$$i_p = KE(1 + m \sin \omega_L t) \sin \omega_H t \text{ (positive values only)} \tag{3}$$

Eq. (3) of course, gives both positive and negative values, and the problem is to obtain from it an equation which will give only positive lobes as indicated in Fig. 3.

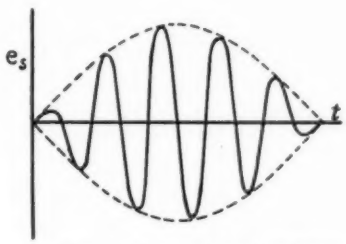

Fig. 2. Form of grid voltage wave.

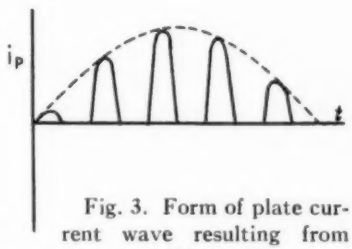

Fig. 3. Form of plate current wave resulting from linear rectification.

Let the required equation be the series,

$$\begin{aligned} i_p = {} & b_0 + a_1 \sin x + a_2 \sin 2x + a_3 \sin 3x + \cdots \\ & + b_1 \cos x + b_2 \cos 2x + b_3 \cos 3x + \cdots \end{aligned} \tag{4}$$

where $x = \omega_L t$.

Rewriting Eq. (3),

$$i_p = KE(1 + m \sin x) \sin qx \text{ (positive values only)} \tag{3′}$$

where

$$q = \frac{\omega_H}{\omega_L}$$

To find the coefficient of the zth sine harmonic, multiply Eq. (4) by $\sin zxdx$ and integrate.

$$\int_0^{2\pi} i_p \sin zxdx = a_z\pi. \tag{5}$$

Combining Eqs. (5) and (3′)

$$a_z a\pi = \int_0^{2\pi} \{KE(1 + m \sin x) \sin qx \text{ (pos. values only)}\} \sin zxdx. \tag{6}$$

The next step is to perform the integration of Eq. (6) taking into account only the positive values of the expression $KE(1+m \sin x) \sin qx$. The cosine coefficients are found in the same way.[3]

There is thus obtained the following series for the rectified plate current of an ideal detector:

$$
\begin{aligned}
i_p = {} & \frac{KE}{\pi} + \frac{KEm}{\pi} \sin \omega_L t + \frac{KE}{2} \sin \omega_H t \\
& - \frac{KEm}{2} \cos (\omega_H + \omega_L)t + \frac{KEm}{2} \cos (\omega_H - \omega_L)t \\
& + \frac{KEm}{\pi} \left[\tfrac{1}{3} \sin (2\omega_H - \omega_L)t - \tfrac{1}{3} \sin (2\omega_H + \omega_L)t \right. \\
& + \frac{1}{3.5} \sin (4\omega_H - \omega_L)t - \frac{1}{3.5} \sin (4\omega_H + \omega_L)t \qquad (7) \\
& \left. + \frac{1}{5.7} \sin (6\omega_H - \omega_L)t - \frac{1}{5.7} \sin (6\omega_H + \omega_L)t + \text{etc.} \right] \\
& + \frac{2KE}{\pi} \left[- \tfrac{1}{3} \cos 2\omega_H t - \frac{1}{3.5} \cos 4\omega_H t \right. \\
& \left. - \frac{1}{5.7} \cos 6\omega_H t - \text{etc.} \right].
\end{aligned}
$$

Conclusions

Concerning ideal detection of the standard signal, these conclusions may be drawn:

(1) A characteristic curve of the form of Fig. 1 does produce ideal or distortionless detection for the standard signal.

(2) The output of the fundamental is linear; i.e., it depends upon E and not upon E^2, and it depends upon m and not upon m^2.

III. Approximation of Ideal Detection

The typical tube characteristic

The static characteristic for ideal detection, the broken line curve of Fig. 1, is impossible of attainment in practice. No tube nor circuit has been devised which has a characteristic bending so sharply. A typical curve for a modern detector tube is shown in Fig. 4.

To approximate ideal operation as closely as possible, the tube is polarized to the knee of the curve. The grid voltage then swings about 0–0′ as an axis, and, to minimize the effect of the curved portion c–c' of the characteristic, the grid swing is made sufficiently large to include the straight line portions

[3] For details of this integration see The University of Texas Bulletin No. 3114, Bureau of Engineering Research.

a–c and c'–b. The operating range a–c–c'–b of the characteristic of Fig. 4 is thus made to resemble the characteristic of Fig. 1 except for the fact that the two straight lines a–c and c'–b of Fig. 4 are connected by a curve c–c' instead of intersecting at a point as in Fig. 1.

The problem

The entire static characteristic curve, using the axis of grid polarization as the zero axis, can be represented by a power series

$$i_p = a_0 + a_1 e_s + a_2 e_s^2 + a_3 e_s^3 + a_4 e_s^4 + \cdots \tag{8}$$

where the series is to be continued until the desired accuracy of representation of the curve is obtained.

Now, having an empirical equation such as (8), the problem is to solve for the direct and audio-frequency current output of a detector tube when a large signal is impressed on the tube.

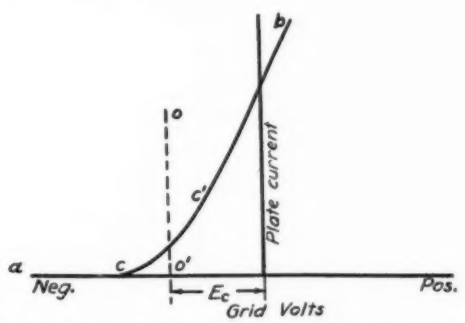

Fig. 4. Typical vacuum tube static characteristic.

Method of solution

For the present, the following limitations are imposed: first, that no grid current flows; second, that the external impedance of the plate circuit is negligible. The equation of the $i_p - e_g$ curve is taken to be of the form (8), where the equation has been set up with the axis of grid polarization as its zero axis. The input to the tube is taken to be the standard signal. The audio-frequency current output is to be expressed as a trigonometric series with the audio-frequency as the fundamental frequency of the series.

For reference, the two known equations and the desired equation are written below:

$$i_p = a_0 + a_1 e_s + a_2 e_s^2 + a_3 e_s^3 + \cdots \tag{8}$$

$$e_s = E(1 + m \sin \omega_L t) \sin \omega_H t. \tag{1}$$

$$\begin{aligned} i_p = {} & c_0 + b_1 \sin \omega_L t + b_2 \sin 2\omega_L t + b_3 \sin 3\omega_L t \\ & + c_1 \cos \omega_L t + c_2 \cos 2\omega_L t + c_3 \cos 3\omega_L t \cdots \\ & + \text{terms of high frequency}. \end{aligned} \tag{9}$$

Eq. (8) is the known tube characteristic; Eq. (1) is the known input voltage;

Eq. (9) is the desired form of the expression for the output current where the coefficients c_0, b_1, b_2, c_1, c_2, $\cdots$ etc. are to be determined.

The general term of the right-hand member of Eq. (8) is $a_n e_s^n$ and, from this point on, the general term will be dealt with. If the value of e_s as given by Eq. (1) is substituted into Eq. (8) the general term becomes:

$$a_n E^n (1 + m \sin \omega_L t)^n \sin^n \omega_H t. \tag{10}$$

By means of the binomial theorem, this term (10) may be expanded, and there results the sub-general term

$$a_n E^n \frac{n!}{r!(n-r)!} m^r \sin^r \omega_L t \sin^n \omega_H t. \tag{11}$$

where r has successively all the integral values, zero to n inclusive.

The trigonometric term $\sin^r \omega_L t$ can be expressed as

$$\begin{aligned} \sin^r \omega_L t = {} & c_{1r} \sin \omega_L t + c_{3r} \sin 3\omega_L t \\ & + c_{5r} \sin 5\omega_L t + \cdots + c_{rr} \sin r\omega_L t \end{aligned} \tag{12}$$

when r is an odd integer;[4] and in the form

$$\begin{aligned} \sin^r \omega_L t = {} & d_{0r} + d_{2r} \cos 2\omega_L t \\ & + d_{4r} \cos 4\omega_L t + \cdots + d_{rr} \cos r\omega_L t \end{aligned} \tag{13}$$

when r is an even integer.[4]

Similarly,

$$\begin{aligned} \sin^n \omega_H t = {} & c_{1n} \sin \omega_H t \\ & + c_{3n} \sin 3\omega_H t + \cdots + c_{nn} \sin n\omega_H t \end{aligned} \tag{14}$$

when n is odd;

$$\begin{aligned} \sin^n \omega_H t = {} & d_{0n} + d_{2n} \cos 2\omega_H t \\ & + d_{4n} \cos 4\omega_H t + \cdots + d_{nn} \cos n\omega_H t \end{aligned} \tag{15}$$

when n is even.

Substitute expansions (12), (13), (14), and (15) into the sub-general term (11):

If r is odd and n is odd, term (11) becomes

---

[4] The coefficient C_{pr} is given by:

$$c_{pr} = (-1)^{(p-1)/2} \frac{1}{2^{r-1}} \frac{r!}{\left(\frac{r-p}{2}\right)!\left(\frac{r+p}{2}\right)!}.$$

Similarly,

$$d_{pr} = (-1)^{p/2} \frac{1}{2^{r-1}} \frac{r!}{\left(\frac{r-p}{2}\right)!\left(\frac{r+p}{2}\right)!};$$

$$d_{0r} = \frac{1}{2^r} \frac{r!}{(r/2)!(r/2)!}.$$

$$a_n E^n \frac{n!m^r}{r!(n-r)!}[c_{1r}\sin\omega_L t + c_{3r}\sin 3\omega_L t + \cdots + c_{rr}\sin r\omega_L t][c_{1n}\sin\omega_H t + c_{3n}\sin 3\omega_H t + \cdots + c_{nn}\sin n\omega_H t]. \tag{16}$$

If r is even and n is odd, term (11) becomes

$$a_n E^n \frac{n!m^r}{r!(n-r)!}[d_{0r} + d_{2r}\cos 2\omega_L t + \cdots + d_{rr}\cos r\omega_L t][c_{1n}\sin\omega_H t + c_{3n}\sin 3\omega_H t + \cdots + c_{nn}\sin n\omega_H t]. \tag{17}$$

If r is odd and n is even, term (11) becomes

$$a_n E^n \frac{n!m^r}{r!(n-r)!}[c_{1r}\sin\omega_L t + c_{3r}\sin 3\omega_L t + \cdots + c_{rr}\sin r\omega_L t][d_{0n} + d_{2n}\cos 2\omega_H t + \cdots + d_{nn}\cos n\omega_H t]. \tag{18}$$

If r is even and n is even,

$$a_n E^n \frac{n!m^r}{r!(n-r)!}[d_{0r} + d_{2r}\cos 2\omega_L t + \cdots + d_{rr}\cos r\omega_L t][d_{0n} + d_{2n}\cos 2\omega_H t + \cdots + d_{nn}\cos n\omega_H t]. \tag{19}$$

There are, then, four possible expansions of the sub-general form (11); these are the expressions (16), (17), (18), and (19). Since n and r are restricted only to positive integers and not to either odd or even integers, all of the above four forms will occur.

The following three trigonometric formulas will be used in performing the multiplications indicated in expressions (16), (17), (18), and (19):

$$\sin x \sin y = -\tfrac{1}{2}\cos(x+y) + \tfrac{1}{2}\cos(x-y); \tag{20}$$

$$\cos x \cos y = \tfrac{1}{2}\cos(x+y) + \tfrac{1}{2}\cos(x-y); \tag{21}$$

$$\sin x \cos y = \tfrac{1}{2}\sin(x+y) + \tfrac{1}{2}\sin(x-y). \tag{22}$$

Referring now to form (16), the performance of the indicated multiplication will result in terms of the type:

$$c_{vn}c_{ur}\sin u\omega_L t \sin v\omega_H t$$

where u and v designate the general terms of the two series. Eq. (20) shows this product to result in terms of sum and difference frequencies. The summation frequencies are of course higher than the carrier frequency, and the difference frequencies may safely be said to be much higher than the audio frequencies. The lowest difference frequency given by expression (16) is

$\cos(\omega_H - r\omega_L)t$. Referring to term (11), r is limited by n, and the lowest difference frequency will depend upon the highest value of n in the power series for the tube characteristic. This lowest difference frequency is, therefore, $\cos(\omega_H - n\omega_L)t$.

Form (16), then, produces no terms of audio-frequency. The same is true of form (17). The conclusion may be drawn that the terms of the power series having odd exponents do not affect the audio frequency.

The general forms (18) and (19) can be treated in the same way, and it will be found that audio-frequency results only from the products

$$(c_{1r} \sin \omega_L t + c_{3r} \sin 3\omega_L t + \cdots + c_{rr} \sin r\omega_L t) d_{0n} \tag{23}$$

and

$$(d_{0r} + d_{2r} \cos 2\omega_L t + \cdots + d_{rr} \cos r\omega_L t) d_{0n} \tag{24}$$

where n must be even, where, in expression (23), r has all odd integral values from zero to n, and where, in expression (24), r has all even integral values from zero to n inclusive.

From these two products, it is easy to pick out the coefficient of any particular harmonic of the audio frequency; e.g., the coefficient of the third sine harmonic is $d_{0n}c_{3r}$, and the coefficient of the zth sine harmonic is $d_{0n}c_{zr}$. But the product (23) is only a typical form. The total coefficient of $\sin z\omega_L t$ is given by

$$\sum_{r=z}^{r=(n-1)} \left(a_n E^r \frac{n!m^r}{r!(n-r)!} d_{0n} c_{rz} \right)$$

where the summation is for the odd values of r only, and where c_{zr} is a function of r.

The total coefficient of $\cos z\omega t$ is obtained from

$$\sum_{r=z}^{r=n} \left(a_n E^n \frac{n!m^r}{r!(n-r)!} d_{0n} d_{zr} \right).$$

General formulas

From the above summations, we can write general formulas for the coefficients of the harmonics of the audio-frequency current.

The general term $a_n e_s{}^n$ of the power series for the static characteristic makes the following contribution to the coefficient of the zth sine harmonic of the audio-frequency:

$$\begin{aligned} a_n E^n \frac{1}{2^{n-1}} \left[\frac{n!}{\left[\frac{n}{2}\right]} \right]^2 (-1)^{(z-1)/2} \Bigg[& \frac{(m/2)^z}{(n-z)!0!z!} + \frac{(m/2)^{z+2}}{(n-z-2)!3!(z+1)!} \\ & + \frac{(m/2)^{z+4}}{(n-z-4)!2!(z+2)!} + \frac{(m/2)^{z+6}}{(n-z-6)!3!(z+3)!} \\ & + \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ & + \frac{(m/2)^{n-3}}{3!\left(\frac{n-3-z}{2}\right)!\left(\frac{n-3+z}{2}\right)!} + \frac{(m/2)^{n-1}}{1!\left(\frac{n-1-z}{2}\right)!\left(\frac{n-1+z}{2}\right)!} \Bigg]. \end{aligned} \tag{25}$$

In the above formula: n must be an even integer greater that z; z must be an odd integer; n is to be less than $w_H/2w_L$.

The general term $a_n e_s^n$ of the power series for the static characteristic makes the following contribution to the coefficient of the zth cosine harmonic:

$$a_n E^n \frac{1}{2^{n-1}}\left[\frac{n'!}{\left[\frac{n}{2}\right]!}\right]^2(-1)^{z/2}\left[\frac{(m/2)^z}{(n-z)!0!z!}+\frac{(m/2)^{z+2}}{(n-z-2)!1!(z+1)!}\right.$$

$$\frac{(m/2)^{z+4}}{(n-z-4)!2!(z+2)!}+\frac{(m/2)^{z+5}}{(n-z-6)!3!(z\pm 3)!} \qquad (26)$$

$$+\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\left.+\frac{(m/2)^{n-2}}{2!\left(\frac{n-2-z}{2}\right)!\left(\frac{n-2+z}{2}\right)!}+\frac{(m/2)^{n-0}}{0!\left(\frac{n-0-z}{2}\right)!\left(\frac{n-0+z}{2}\right)!}\right]$$

In the above formula: n must be an even integer greater than or equal to z; z must be an even integer; n is to be less than $w_H/2w_L$. Formula (26) differs from formula (25) only in the term which determines the sign, and in the fact that the integer z must be even rather than odd.

The general term $a_n e_s^n$ of the power series for the static characteristic makes the following contribution to the direct current output of the tube:

$$a_n E^n \frac{1}{2^n}\left[\frac{n!}{\left[\frac{n}{2}\right]!}\right]^2\left[\frac{1}{n!}+\frac{(m/2)^2}{(n-2)!1!1!}+\frac{(m/2)^4}{(n-4)!2!2!}\right.$$

$$+\frac{(m/2)^6}{(n-6)!3!3!}+\cdots \qquad (27)$$

$$+\cdots\cdots\cdots$$

$$+\frac{(m/2)^{n-4}}{4!\left(\frac{n-4}{2}\right)!\left(\frac{n-4}{2}\right)!}$$

$$\left.+\frac{(m/2)^{n-2}}{2!\left(\frac{n-2}{2}\right)!\left(\frac{n-2}{2}\right)!}+\frac{(m/2)^n}{0!\left(\frac{n}{2}\right)!\left(\frac{n}{2}\right)!}\right].$$

In formula (27) n must be an even integer.

The coefficients for direct current and for the fundamental audio-frequency and its low harmonics

Using the above general formulas, it is possible to write:

The direct current output of the tube:

$$
\begin{aligned}
&a_0 \\
&+ a_2E^2[0.5 + 0.25m^2] \\
&+ a_4E^4[0.375 + 1.125m^2 + 0.140625m^4] \\
&+ a_6E^6[0.3125 + 2.34375m^2 + 1.757812m^4 + 0.90765625m^6] \\
&+ a_8E^8[0.2734375 + 3.828125m^2 + 7.177734m^4 + 2.392578m^6 \\
&\qquad + 0.07476807m^8] \\
&+ a_{10}E^{10}[0.2460937 + 5.537109m^2 + 19.37988m^4 + 16.14990m^6 \\
&\qquad + 3.028107m^8 + 0.06056213m^{10}] \\
&+ \cdots
\end{aligned}
\tag{28}
$$

The fundamental output (i.e., the coefficient of sin $\omega_L t$):

$$
\begin{aligned}
&a_2E^2[0.25m] \\
&+ a_4E^4[1.5m + 1.125m^3] \\
&+ a_6E^6[1.875m + 4.6875m^3 + 1.171875m^5] \\
&+ a_8E^8[2.1875m + 11.484375m^3 + 9.570312m^5 + 1.196289m^7] \\
&+ a_{10}E^{10}[2.460937m + 22.14844m^3 + 38.75987m^5 + 16.14990m^7 \\
&\qquad + 1.211243m^9] \\
&+ \cdots
\end{aligned}
\tag{29}
$$

The second harmonic output (coefficient of cos $2\omega_L t$):

$$
\begin{aligned}
&- a_2E^2[0.25m^2] \\
&- a_4E^4[1.125m^2 + 0.1875m^4] \\
&- a_6E^6[2.34375m^2 + 2.34375m^4 + 0.1464844m^6] \\
&- a_8E^8[3.828125m^2 + 9.570312m^4 + 3.588867m^6 + 0.1196289m^8] \\
&- a_{10}E^{10}[5.537109m^2 + 25.83984m^4 + 24.22485m^6 + 4.844971m^8 \\
&\qquad + 0.1009369m^{10}] \\
&- \cdots
\end{aligned}
\tag{30}
$$

The third harmonic output (coefficient of sin $3\omega_L t$):

$$
\begin{aligned}
&- a_4E^4[0.375m^3] \\
&- a_6E^6[1.5625m^3 + 0.5859375m^5] \\
&- a_8E^8[3.828125m^3 + 4.785156m^5 + 0.7177734m^7] \\
&- a_{10}E^{10}[7.382812m^3 + 19.37988m^59.68994m^7 + 0.8074949m^9] \\
&- \cdots
\end{aligned}
\tag{31}
$$

It will be found that those terms in expressions (28), (29), (30), and (31)

which involve the parameter a_2, check with the corresponding terms derived from small signal theory where parameters higher than a_2 are not included.

Effect of grid polarization

The above discussion was based on the condition that the equation of the $i_p - e_g$ curve had as its zero axis the axis of polarization. A change in the grid bias can be taken into account by shifting the zero axis of the power series which represents the $i_p - e_g$ characteristic.

Conclusions

From Eqs. (23), (24), (25), and (26) and their derivations, it is possible to draw some interesting conclusions:

1. Only the even exponent members of the power series produce d.c. and the audible frequencies.
2. The nth term of the power series (i.e., $a_n e_s^n$) can introduce no harmonic higher than the nth but affects all harmonics lower than this.
3. The sine harmonics are all odd.
4. The cosine harmonics are all even.
5. As the signal strength increases, the importance of the higher terms of the power series increases.
6. As the percentage modulation increases, the importance of the higher terms of the power series increases.

Extension to include pure resistance loading in grid and plate circuits

The method of analysis given in Section III has been based upon the condition that the relation between the voltage impressed upon the grid circuit and the current flowing in the plate circuit is defined by a single valued curve the empirical equation of which can be expressed in the form of a power series. This is true when the detector tube alone is considered. It is also true when the impedances combined with the tube are pure resistances. However, when the impedances are functions of frequency, the relation between plate current and varying grid voltage is defined, not by a single valued curve, but by a series of closed curves.

In order to apply the previous analysis to the case where the tube is used in connection with pure resistances, it is only necessary to use as a basis the static characteristic of the detection circuit rather than the characteristic of the tube alone. The detection circuit may be set up, and its static characteristic can then be taken by the ordinary d.c. methods. Or it is possible to use the static characteristics of the tube alone and, by graphical means, derive from them the static characteristic of the detection circuit for any particular circuit constants which may be chosen.[5]

IV. Power Series Analysis of Static Characteristic

Method of simultaneous equations

The previous analysis of detector performance has used as a basis Eq.

[5] The University of Texas Bulletin No. 3114, Bureau of Engineering Research.

(26), an empirical relation between plate current and signal voltage expressed in the form of a power series. If the static characteristic represented by such a series is simple, or if the range covered is small, the parameters can be solved for by the method of simultaneous equations. For example, if four terms of the series are sufficient, it is possible to choose four points on the curve, solve four simultaneous equations, and obtain the constants a_0, a_1, a_2, and a_3. The graph of the empirical equation passes through the four chosen points.

However, the characteristics to be analyzed can seldom be represented over any large range by a four-term series. If a close fit is desired, and if the series is extended to say $a_{10}e^{10}$ in order to obtain this close fit, then it becomes necessary to solve eleven simultaneous equations. Further, though the empirical equation so found must pass through the eleven chosen points, in between these points its graph may not fit the original curve at all.

Fourier series method

The following method avoids the above difficulties.[6]

Suppose that the region E_0-0-E_0' of the curve is to be represented by a power series, the zero axis of the equation to be the axis of grid swing, which is located at the center of the distance E_0-E_0'. There is to be obtained an equation of the form

$$i_p = \alpha_0 + \alpha_1 e_s + \alpha_2 e_s^2 + \alpha_3 e_s^3 + \cdots$$

where $e_s = e_0 - E_c$. Express the variable e_s as the following function of the arbitrary variable x:

$$e_s = E_0 \sin x. \tag{32}$$

Replot the curve replacing e_s by the new variable x. There is obtained a periodic function of x. If this periodic function is analyzed by the Fourier series method, there results the equation,

$$\begin{aligned} i_p = {} & c_0 + b_1 \sin x + b_3 \sin 3x + b_5 \sin 5x + \cdots \\ & + c_2 \cos 2x + c_4 \cos 4x + c_6 \cos 6x + \cdots . \end{aligned} \tag{33}$$

The even sine and the odd cosine terms of the series are excluded because the curve is symmetrical about a vertical axis through the point $x = 90°$.

Now $\sin 3x = 3 \sin x - 4 \sin^3 x$; $\sin 5x = 5 \sin x - 20 \sin^3 x + 16 \sin^5 x$; etc. But, from Eq. (32),

$$\sin x = e_s/E_0. \tag{34}$$

[6] De La Vallee Poussin, L'Approximation des Fonctions, pp. 63, 64. The writer was referred to this book after his development of this method.

Therefore,

$$\sin 3x = \frac{3e_s}{E_0} - \frac{4e_s^3}{E_0^3};$$

$$\sin 5x = \frac{5e_s}{E_0} - \frac{20e_s^3}{E_0^3} + \frac{16e_s^5}{E_0^5};$$

$$\cos 2x = 1 - \frac{2e_s^2}{E_0^2};$$

$$\cos 4x = 1 - \frac{8e_s^2}{E_0^2} + \frac{8e_s^4}{E_0^4}.$$

And so, using the above relations, the Fourier series of Eq. (33) can readily be transformed into a power series expressing i_p as a function of e_s over the range E_0-0-E_0'. It will be noted that, if the Fourier series ends with the angle $11x$, the power series ends with the term e_s^{11}; if the Fourier series ends with the angle $12x$, the power series ends with the term e_s^{12}; and so on.

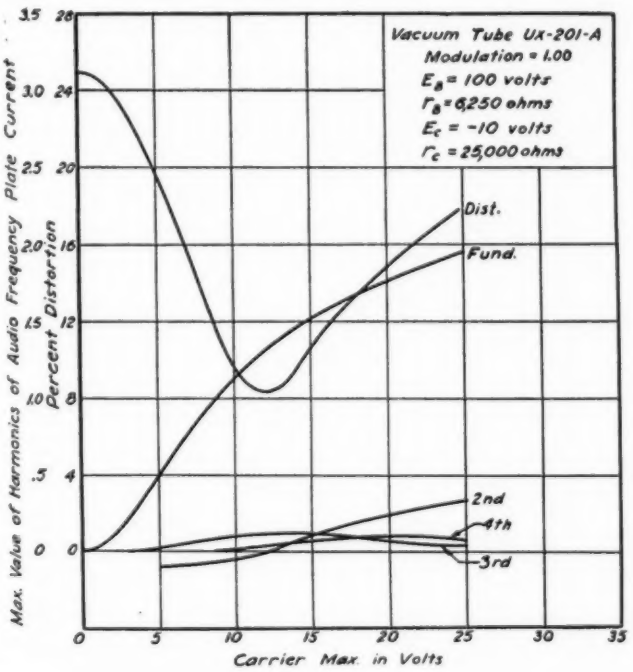

Fig. 5.

Advantages of Fourier series method

The above method reduces the power series analysis of a given curve to the Fourier series analysis. It thus avoids simultaneous equations and the raising of numbers to the higher powers. Moreover—and here lies the chief advantage—it is not necessary to obtain the power series before the closeness of fit can be tested. When Eq. (33) has been obtained, the closeness of fit can

easily be tested without any need of raising e_s to a series of high exponents only to find that the series is not long enough. If it develops that Eq. (33) does not fit closely enough, the Fourier series can readily be extended, and none of the previous work must be discarded. When the Fourier series is long enough to be sufficiently accurate, it can be transformed into the power series, and the analysis is complete with no need for testing or extending the power series.

V. Analysis of Detection Performance of UX201-A Vacuum Tube

Comparison with experimental results

Using the formulas developed, detection performance has been calculated for a UX201-A tube having external resistances in the grid and plate circuits. The results are plotted in Fig. 5, 6, 7, and 8. Ballantine[7] has obtained

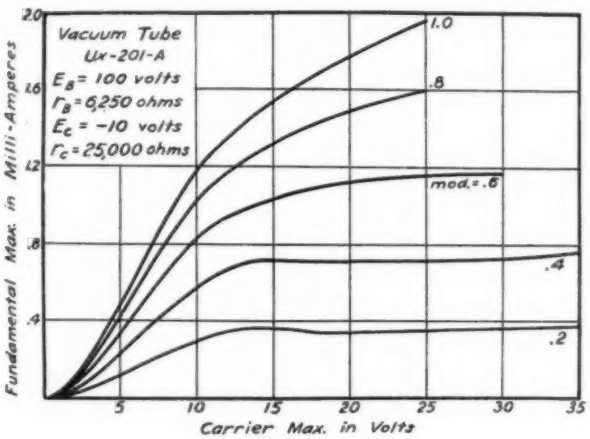

Fig. 6.

these curves experimentally. It is interesting to compare the two sets of curves qualitatively. Since different circuit constants were used in each case, they cannot be compared quantitatively. The distortion curves all show the same pronounced dip at some critical value of signal strength. Further, this critical value of E is little affected by the percentage modulation so that the best operating point of the circuit is the same for all modulations. In both the calculated and the experimental curves, the best operating point is reached after grid current begins. Ballantine mentions that the second harmonic, as obtained experimentally, decreases as E increases becoming zero near that value of E which makes distortion a minimum. At this point the second harmonic reverses phase and then increases as E increases. The second harmonic curve of Fig. 5 is in agreement with Ballantine's results.

[7] S. Ballantine, Detection at High Signal Voltages. Part I—"Plate Rectification with the High Vacuum Triode", Institute of Radio Engineers, Proc. **17**, 1153 (1929).

For low modulations, Ballantine obtained curves showing that the fundamental output first increased with E, reached a maximum at some value of E greater than that which gave minimum distortion, and then decreased sharply with a further increase in E. For low modulations, the calculated curves

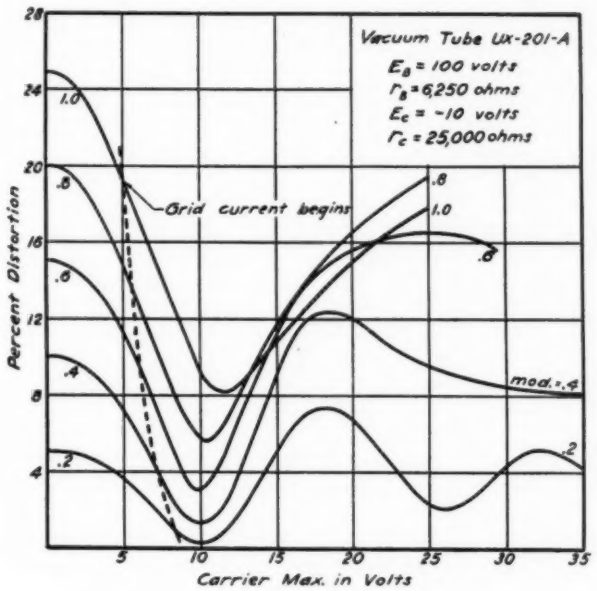

Fig. 7.

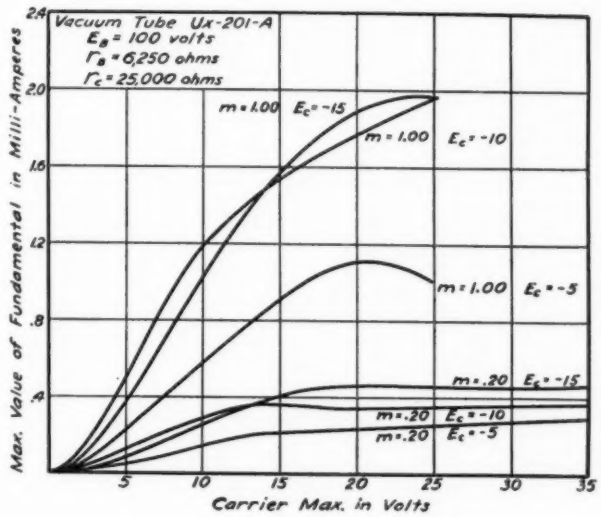
Vacuum Tube UX-201-A
E_b = 100 volts
r_a = 6,250 ohms
r_c = 25,000 ohms
Max. Value of Fundamental in Milli-Amperes
Carrier Max. in Volts
m = 1.00 E_c = -15
m = 1.00 E_c = -10
m = 1.00 E_c = -5
m = .20 E_c = -15
m = .20 E_c = -10
m = .20 E_c = -5

Fig. 8.

of Fig. 6 show a slight decrease in fundamental output when E becomes greater than 15 volts, but nothing like the decrease Ballantine obtained. The difference is probably due to the difference in circuit constants. Refer to the curve of Fig. 8 for $m = 1.00$, $E_c = -5$ volts. Apparently the dip in the curve

occurs when the polarization point is shifted away from the knee of the static characteristic.

Effect of the grid bias

The curves of Figs. 8 and 9 have been plotted to show the effect of variation of E_c. With the exception of E_c, the circuit constants are the same as before. The modulation is taken first as 1.00 then as 0.20, and E_c is allowed to take the values -15 volts, -10 volts, and -5 volts.

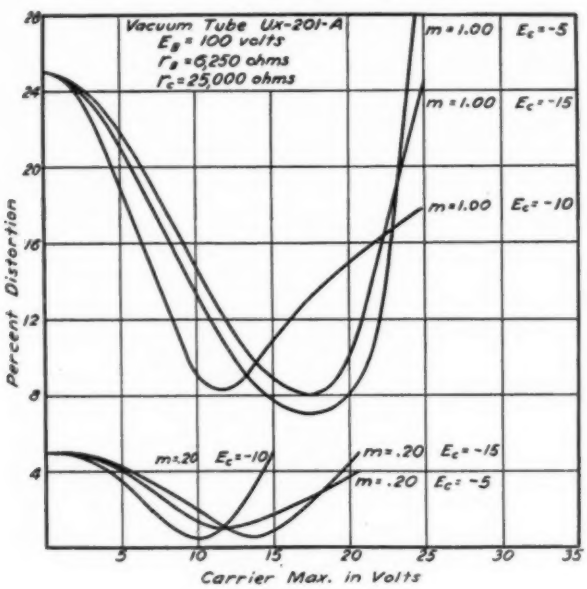

Fig. 9.

The principal effect of moving the polarization point away from the knee of the characteristic appears to be to increase the value of E which gives minimum distortion. Thus:

for $E_c = -10$ volts (knee of the characteristic), minimum distortion is obtained for $E = 11$ volts;

for $E_c = -5$ volts, minimum distortion is obtained for $E = 17.5$ volts;

for $E_c = -15$ volts, minimum distortion is obtained for $E = 17.5$ volts.

The variation in E_c does not greatly affect the actual value of the minimum distortion; Fig. 9 indicates that, for 100 percent modulation, it is possible to obtain as low as 8 percent distortion for all three values of E_c. After the minimum point is passed, the curves for $E_c = -5$ volts and $E_c = -15$ volts show the distortion to increase very rapidly with E, much more rapidly than for the case $E_c = -10$ volts.

Table of Symbols

e_0—total voltage in the grid circuit.

$$e_0 = e_s + E_c.$$

e_s—signal voltage. Throughout this paper the signal is assumed to be the standard,

$$e_s = E(1 + m \sin \omega_L t) \sin \omega_H t.$$

E_c—d.c. grid bias.
E—amplitude of the carrier wave.
m—percent modulation.
ω_L—$2\pi f_L$, where f_L is the audio frequency.
ω_H—$2\pi f_H$, where f_H is the carrier frequency.
r_c—external grid resistance.
r_B—external plate resistance.

Bibliography

S. Ballantine, Detection at High Signal Voltages. Part I—"Plate Rectification with the High Vacuum Triode", Institute of Radio Engineers, Proc. **17,** 1153 (1929).

G. L. Beers, "Power Detection for Broadcast Reception", Electric Journal **26,** 413 (1929).

G. Breit, "The Calculation of Detecting and Amplifying Properties of an Electron Tube from Its Static Characteristics", Phys. Rev. **13,** 387 (1920).

J. R. Carson, "A Theoretical Study of the Three Element Vacuum Tube", Institute of Radio Engineers, Proc. **7,** 187 (1919).

E. L. Chaffee and G. H. Browning, "A Theoretical and Experimental Investigation of Detection for Small Signals", Institute of Radio Engineers, Proc. **15,** 113 (1927).

F. B. Lewellyn, "Operation of Thermionic Vacuum Tube Circuits", Bell System Technical Journal **5,** 432 (1926).

W. B. Medlam, "Improving Detector Efficiency", Wireless World **24,** 524 (1929).

W. I. G. Page, "The Valve as an Anode Bend Detector", Wireless World **24** 279 (1929).

F. E. Terman and N. R. Morgan, "Some Properties of Grid Leak Power Detection", Institute of Radio Engineers, Proc. **18,** 2160 (1930).

NEW THEORY AND DESIGN OF WAVE FILTERS*

By W. Cauer
Göttingen, Germany

(Received November 12, 1931)

Abstract

A new mathematical theory of wave filters gives the following results: It is possible to realize as constant an image impedance as desired in the transmitting bands. In the new theory there is no mathematical difference in the consideration of proper attenuation or proper image impedance characteristics. In other words, the same attack applies in both cases. All known wave filters as far as they are symmetrical four-terminal networks or equivalent to such networks in series with a transformer, are contained in the filters of the new theory which comprises all possible filters of this kind. The new theory has been carried through practically to find out the most economical filter for any practical purpose. This results in many cases in solving a given practical question already solved by known methods with considerable saving of material. The chief aim of this paper is to show the relation between the new theory and the older design methods due to O. J. Zobel and other members of the Bell Laboratories.

SINCE O. J. Zobel published his paper "Theory and Design of Uniform and Composite Electric Wave Filters"[1] beginning with the words "The electric wave-filter as regards its general transmission characteristics and its extremely important role in communication systems, is well known," his own theory and methods of designing filters have also become well known and have come into general use.[2] Therefore a short review of his methods should suffice.

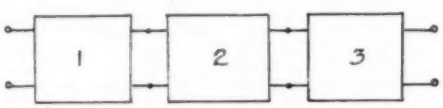

Fig. 1. Filter chain.

The filters used by Zobel are generalizations of the filter chains with identical sections, invented independently by G. A. Campbell and K. W. Wagner. The several sections (1, 2, 3 in Fig. 1) may be general passive[3] four-terminal networks (symmetrical or unsymmetrical) but such that a certain condition of matching is satisfied between any two successive sections. To understand this precisely we define the following three parameters (functions of frequency) which characterise completely the exterior properties of any four-

* The first short communication of the new theory was made in a lecture given November, 27, 1928, in Göttingen.

[1] The Bell System Technical Journal (1923).

[2] See, for instance, the book of K. S. Johnson "Transmission Circuits for Telephonic Communication", and the book of T. E. Shea "Transmission Networks and Wave Filters".

[3] "Passive" means that no electromotive forces or tubes or negative impedance elements are contained in the network.

terminal network F. Image impedances are defined as those impedances W_1, W_2 which are uniquely determined by the following requirement: F with W_2 connected across the ends 2, 2 (Fig. 2) gives a driving point impedance W_1 at the ends 1, 1. And at the same time F with W_1 connected across the ends 1, 1 gives a driving point impedance W_2 at the ends 2, 2.

The propagation constant Γ is defined as

$$\Gamma = \log_e J_1/J_2 = A_1 + iB_1$$

where J_1 is the complex input current and J_2 is the complex output current in the case where the four-terminal network F is connected with the image impedances W_1 and W_2 as shown in Fig. 2, an electromotive force being assumed in series with the impedance W_1.

The condition of matching mentioned above may be stated as follows: The image impedances at the connected ends of two successive sections must be equal. It is obvious that in any chain which satisfies this condition the image impedances of the whole chain are identical with the image impedances at the free ends of the first and last section of the chain, and the propagation constant of the whole chain is the sum of the propagation constants of all sections.

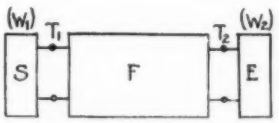

Fig. 2. Filter connected with sending and receiving apparatus.

It follows that it is not difficult to calculate the behavior of such a chain structure as a function of frequency if the frequency characteristics of the single sections are known. By combining several comparatively simple sections in the manner just described it becomes possible to realize attenuation characteristics A_1 which are much better than those obtainable by single sections with a restricted number, say 6, circuit elements. As was stated above, the image impedances of the whole chain are determined by the end sections. There are several devices developed to obtain proper image impedances by suitable end circuits, the best known of which is perhaps Zobel's M-type.[4]

What we really need in a filter is a working attenuation A (total attenuation) which is nearly zero in prescribed frequency ranges and greater than a given constant in the main part of the complementary frequency ranges. The total attenuation may be defined as $A = \log_e (|J'/J_2|)$, J' denoting the current which flows in the receiving apparatus if sending apparatus S and receiving apparatus E are immediately connected, J_2 denoting the corresponding current if the wave filter F is put between these two pieces of apparatus (Fig. 2). Only if S and E have impedances W_1 and W_2 equal to the image impedances W_1 and W_2 of the corresponding ends of the filter F, will the total attenuation be equal to the "attenuation constant" A_1, as defined before.

[4] See also papers of Feldtkeller and Strecker for instance Elektrische Nachtrichtentechnik, June, 1927, and H. W. Bode, B.S.T.J., October, 1930.

Without restriction on generality (by using a suitable transformer) we may assume that those impedances of S and E are practically the same real constant R in the transmitting range (or transmitting ranges). The effect of not matching image impedances ($W_1 \neq R$, $W_2 \neq R$) on the total attenuation is such that additional terms must be added to the attenuation constant A_1 which depend on W_1, W_2 and R. It is known that, taking a comparatively simple four-terminal section, it is possible to have $A_1 > 0$ in the prescribed attenuation ranges and $A_1 = 0$ in the complementary ranges. Hence by taking a chain with a sufficiently large number of such sections it is obvious that we succeed in a nearly ideal A_1-frequency characteristic. Because A_1 is usually numerically the chief part of the total attenuation A, the filters of the kind just mentioned satisfy in a rough way those formulated ideal requirements. Mathematically the ideal A_1 or the known chain structures solve only half of the problem by making the total attenuation sufficiently high in the attenuation bands, but not guaranteeing a nearly zero attenuation in the transmitting ranges. For satisfying the latter condition with a high degree of accuracy, we need filters with nearly constant image impedances ($=R$) in the transmitting ranges. But there is another more important practical reason for this image impedance requirement; in order that echo effects in connecting long lines with filters may be low the image impedance of the filter should be as constant as possible in the transmitting ranges.

The Zobel method of designing wave filters consists essentially in combining suitably matched sections to a chain structure. On the one hand it becomes possible by this method to calculate practically the attenuation, image impedance and total attenuation characteristics of quite complicated four-terminal structures by tabulating only the characteristics of a series of comparatively simple four-terminal networks. For appreciating the advantages of the new theory it must be considered that up to the present:

(1). Although it has been shown that by taking chains with a sufficiently high number of sections, it is possible to satisfy any given requirement for the attenuation constant A_1, there is no certainty whether in this way the most economical solution with the minimum number of circuit elements is obtained.

(2). Although the image impedances were improved by suitable termination, (a) it has not been shown mathematically up to the present, that it is possible to realize as constant an image impedance as is desired in the transmitting bands; (b) moreover not only the image impedance but also the attenuation is influenced by the terminating circuits. Thus by methods of suitable termination an image impedance requirement can not be satisfied without changing the characteristic of the attenuation constant A_1; (c) suitable termination does not necessarily give the most economical solution of an image impedance requirement but sometimes contains superfluous circuit elements.

Much as we are indebted to the numerous researches of the electrical engineers in the field of filter-networks, a general theory of wave-filters has not existed until now.

In the new theory which I shall outline the essential points are: (1). The joint treatment of the circuits which are equivalent for all frequencies; (2). The conception of those variables which characterize a $2n$-terminal network as analytic functions of a complex variable λ ($\lambda = i\omega$, ω being the angular velocity in radians per second); (3). The exact study of the properties of these complex functions arising in such networks.

This method of considering circuits is useful and the most general for the solution of all questions of communication engineering which demand that circuits of any desired dependence in frequency be realized. Besides wave-filters, further well-known examples are simulating, attenuation and phase correction circuits. The corresponding pure mathematical question always involves the solution of a certain interpolation problem in which only the special kinds of functions arising in (3) are admitted as interpolating functions.

In this paper an application is made of these quite general principles to symmetrical four-terminal networks and more particularly to symmetrical electric wave filters.

A four-terminal network is called symmetrical, if the external behavior does not change by changing the output and input terminals, or, what is the same thing, if $W_1 = W_2$.

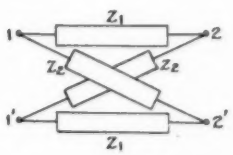

Fig. 3. Canonical symmetrical four-terminal network.

We start with the following theorem: *Any symmetrical four-terminal network is physically equivalent* to a bridge type or lattice network (one section only) in which each pair of opposites branches has the same driving point impedances (Z_1 and Z_2) (see Fig. 3).*

A special case of this theorem was published first by G. A. Campbell in "Physical Theory of Wave Filters".[5] The first general proof of the theorem is to be found in my paper "Über die Variabeln eines passiven Vierpols".[6]

The behavior of a general unsymmetrical four-terminal network is described completely by three functions of frequency, for instance Γ, W_1 and W_2. These three functions are reduced to two in the case of symmetry $W_1 = W_2$. In this case however we shall prefer two other characterizing functions, namely, Z_1 and Z_2 the rational functions of λ which occur in the above stated theorem of equivalence. These impedances Z_1 and Z_2 have the simple physical meaning of driving point impedances in the Wheatstone bridge. We are allowed to neglect the resistances from the standpoint of theory, because transmitting bands with zero attenuation demand zero resistances. The charac-

* *Equivalent* in this paper always means equivalent for all frequencies.

[5] Campbell, Bell Syst. Techn. Journ. (1922).

[6] Cauer, Sitzungsberichte der Preuss. Akademie der Wiss. (1927).

teristic form of a driving point impedance (that is of $Z_1(\lambda)$ and $Z_2(\lambda)$) without resistances is given by a theorem of R. M. Foster.[7]

The necessary and sufficient condition for the realizability of a driving-point impedance function z of frequency $\lambda = i\omega$ without resistances and with a finite number of circuit elements (positive inductances, mutual inductances, capacities) is that the function may be written in the following form:

$$z = H \frac{(\lambda^2 + \omega_1^2)(\lambda^2 + \omega_3^2) \cdots (\lambda^2 + \omega_{2n-1}^2)}{\lambda(\lambda^2 + \omega_2^2) \cdots (\lambda^2 + \omega_{2n-2})} \tag{1}$$

where H is a positive constant and the resonant and the anti-resonant frequencies satisfy the conditions $0 < \omega_1 < \omega_2 < \cdots < \omega_{2n-1} < \omega_{2n-1}$. The limiting cases $\omega_1 = 0$, $\omega_{2n-1} = \infty$ (this means the last factor in the numerator is to be omitted) are allowed.

An equivalent formulation of this theorem which is more convenient from the standpoint of the theory of complex variables is the following: *Any rational function of $\lambda = i\omega$ which satisfies the following conditions and only such a function is physically realizable as a driving point impedance without resistances. The function is regular in the interior of the right λ-half plane and has there a positive real part. The function is real on the real λ-axis and pure imaginary on the imaginary λ-axis (except for poles which are necessarily simple).*

If in the following we use the impedances Z_1 and Z_2 for characterizing a filter network and reduce the design problem of a filter to that of deriving suitable functions Z_1 and Z_2. No restriction is made by this procedure of assuming a bridge type circuit because any symmetrical filter network may be characterized by our Z_1 and Z_2 which are both functions of the form (1). If such functions of the form (1) are given there are several known methods of deriving physical two-terminal circuits realizing these functions as for instance the process of separation in partial fractions or development in continued fractions or a combination of these two processes. It is unnecessary to point out in detail here the processes which correspond to certain ladder structures, because they are given the above mentioned publications of R. M. Foster and myself. The question of finding all other equivalent circuits is treated in my paper "Vierpole".[8] In more accurate and complete form the fundamental group of equivalence of general four-terminal networks is discussed in my paper "Untersuchungen über ein Problem, das drei positiv definite quadratische Formen mit Streckenkomplexen in Beziehung setzt".[9,10] The given principles for deriving equivalent circuits may be applied to all four-terminal networks, not only to symmetrical networks. This means, for instance, that they may be applied to symmetrical filters in series with a transformer. This is of interest in so far as it shows the general application of

[7] Foster, Bell Syst. Techn. Journ. (1924); compare also Cauer, Archiv f. Elektrotechnik, Bd. 705 Dec. (1926).

[8] Elektrische Nachrichtentechnik (1929).

[9] Cauer, Mathematische Annalen (1931).

[10] Another paper is planned to discuss the question of equivalence for four-terminal networks without resistances.

the theory and design method given here to cases for which the sending apparatus S and the receiving apparatus E have different impedance values.

In the next section it will be shown how to determine Z_1 and Z_2 in order to obtain any desired symmetrical wave filters. Afterwards it will be pointed out in how far the known filters are only special cases of the general filters derived by the new theory.

Instead of taking Z_1 and Z_2 themselves it is more convenient for characterizing a filter to take the product Z_1Z_2 and the quotient Z_2/Z_1. For the propagation constant Γ, and therefore the attenuation constant A_1 which is the real part of Γ, depends only on the quotient Z_2/Z_1 according to the equation

$$\Gamma = \log_e \frac{(Z_2/Z_1)^{1/2} + 1}{(Z_2/Z_1)^{1/2} - 1}. \tag{2}$$

On the other hand the image impedance (characteristic impedance) Z depends only on Z_1Z_2 according to the simple relation

$$Z = (Z_1Z_2)^{1/2}. \tag{3}$$

To reduce echo effects Z should approximate a positive constant R in the transmitting ranges, R being the impedance of the given sending apparatus S and the receiving apparatus E. But this follows also mathematically from the attenuation requirement which gives us a requirement for the quotient $(Z_2/Z_1)^{1/2}$. As stated above the total attenuation A depends not only on the filter network (Z_1 and Z_2) itself, but on the resistance R of the connected apparatus as well. According to the definition of A given above, we easily get the formula (see appendix 1)

$$\begin{aligned} A &= \log_e \left| \frac{(1 + Z_1/R)(1 + Z_2/R)}{Z_2/R - Z_1/R} \right| \\ &= \log_e \left| \frac{(1 + z_1)(1 + z_2)}{z_2 - z_1} \right|, \end{aligned} \tag{4}$$

by introducing the reduced impedance functions of the kind (1)

$$z_1 = Z_1/R, \quad z_2 = Z_2/R. \tag{5}$$

By using this notation a more symmetrical mathematical formulation of our technical requirements becomes possible.

We want to have (a) A nearly zero in the given transmitting frequency ranges (these ranges are certain given intervals of the pure imaginary λ-axis), and (b) A nearly infinite in the complementary frequency ranges. The first condition is equivalent to the following condition:

$$\left| \frac{(1 + z_1)(1 + z_2)}{z_2 - z_1} \right| \approx 1$$

in the transmitting ranges. By this requirement the case $z_2 - z_1 = 0$ or $1/z_1 - 1/z_2 = 0$ is excluded. Since according to (1) z_1 and z_2 are pure imaginary,

by taking the square of the absolute value the ideal requirement may be written in the form

$$\frac{(1 + z_1 z_2)^2 - (z_1 + z_2)^2}{-(z_2 - z_1)^2} = 1,$$

or

$$z_1 z_2 = 1 \text{ in the transmitting ranges.} \tag{6}$$

According to (3) and (5) this is the same as $Z = R$. This is also the condition stated above for reducing the echo effects. Because z_1 and z_2 may become infinite or zero in single points only, according to (1), the requirement (b) obviously reduces to

$$z_2/z_1 = 1 \text{ in the attenuation ranges.} \tag{7}$$

This condition is equivalent to $A_1 = \infty$ which follows from the fact that A_1 is the real part of (2) (Γ) considering the definition (5). With $z_1(\lambda)$ the function $1/z_1(\lambda)$ is according to (1) physically realizable as driving point impedance without resistances. Hence the interpolation problem for approximately satisfying the ideal requirements (6) and (7) is mathematically the same for (6) as for (7).

By the method described above of taking a chain with a sufficiently large number of equal sections, each section being a simple symmetrical four-terminal network satisfying the conditions

$$A_1 = 0 \text{ in the transmitting ranges,}$$
$$A_1 > 0 \text{ in the attenuation ranges}$$

it has been shown physically (and of course indirectly mathematically too) that (7) may be satisfied with any desired degree of accuracy by functions z_1 and z_2 which are of the form (1). But, because the interpolation problem of satisfying (6) approximately by rational functions z_1 and z_2 of the type (1) is mathematically the same as the corresponding interpolation problem for (7), the chain method proves, at the same time, the possibility of getting as constant an image impedance Z in the transmitting bands as we may prescribe. This proof gives us also a first method of design of filters with image impedances as accurate as we need. What we have to do is to calculate the function Z_1 and Z_2 of a chain (see below) and then realize (for instance) R^2/Z_1 and Z_2 as the driving point impedances in the canonical bridge type (lattice type) network (Fig. 3). Generally speaking, if we consider a filter with characteristics Z_1 and Z_2, another complementary filter is obtained which attenuates where the first one transmits, and vice versa, and is good for the same impedance R of the sending and receiving apparatus (S and E in Fig. 2), by only changing Z_1 into R^2/Z_1 (z_1 into $1/z_1$), or Z_2 into R^2/Z_2 (z_2 into $1/z_2$). For instance such a transformation converts a band-pass filter into a band-elimination filter.

It is equally obvious how to obtain from a given filter a *reciprocal or in-*

* A general definition of reciprocal $2q$-terminal networks of which the above definition of reciprocal symmetrical filters is a special case, is given in the appendix II.

*verse filter** which has the same transmitting and attenuating ranges and is proper for the same impedance R of the connected apparatus, but has a reciprocal image impedance characteristic (R^2/Z) except for the positive constant factor R^2. We must change Z_1 into R^2/Z_1 (z_1 into $1/z_1$) and *at the same time* Z_2 into R^2/Z_2 (z_2 into $1/z_2$). Of course the approximations of the ideal requirements (6) and (7) are both invariant for such substitution.

The ideal requirements (6) and (7) show that even for a rough approximation of these conditions no zeros or poles of z_2/z_1 are allowed in the attenuation bands, and no zeros or poles of z_1z_2 are allowed for the transmitting bands. But the last condition signifies nothing else than coincidence of the zeros (resonance frequencies) of z_1 with the poles (anti-resonance frequencies) of z_2 in the transmitting ranges, and, simularly, the coincidence of the poles of z_1 and the zeros of z_2 in the same ranges. Furthermore the first condition means the coincidence of the zeros of z_1 with the zeros of z_2 and the poles of z_1 with the poles of z_2 in the attenuation ranges. It follows that $(z_2/z_1)^{1/2}$ has, as function of λ^2, simple zeros and simple poles separate from each other (compare (1)) in the interior of all transmitting ranges and no other zeros and poles. Simularly $(z_1z_2)^{1/2}$ has simple zeros and simple poles which are separate from each other in the interior of all attenuation ranges and no other zeros or poles. Because in the transmitting ranges $z_1z_2>0$ (or Z real or $z_2/z_1<0$) and in the attenuation ranges $z_2/z_1>0$ (or $z_1z_2<0$ or Z pure imaginary), the functions z_1z_2 and z_2/z_1 that are real for pure imaginary λ have necessarily a change of sign at any "cut-off frequency", which is defined to separate a transmitting band from an attenuation band. This means that at each cut-off frequency either z_1 or z_2 has a zero or a pole, but not both together, or, what is the same, $(z_2/z_1)^{1/2}$ and $(z_1z_2)^{1/2}$ have irrational factors $(\lambda^2+\omega_r^2)^{1/2}$ or $(\lambda^2+\omega_r^2)^{-1/2}$. Because zeros and poles in z_1 or z_2 separate each other according to (1), the first factor arises when (and only when) a pole of $(z_2/z_1)^{1/2}$ (or $(z_1z_2)^{1/2}$) precedes or follows and the second factor, if a zero of $(z_2/z_1)^{1/2}$ (or $(z_1z_2)^{1/2}$) precedes or follows. Finally, it results from (1) that $(z_2/z_1)^{1/2}$ (or $(z_1z_2)^{1/2}$) for the frequency $\omega=0$ has a factor λ if a pole follows and a factor $1/\lambda$ if a zero of $(z_2/z_1)^{1/2}$ (or $(z_1z_2)^{1/2}$) follows, but $(z_2/z_1)^{1/2}$ (or $(z_1z_2)^{1/2}$) has neither the factor λ not the factor $1/\lambda$ if its zeros and poles begin above the first cut-off frequency. With the foregoing the characteristic feature of the functions $(z_2/z_1)^{1/2}$ and $(z_1z_2)^{1/2}$ for a wave filter with given transmitting and attenuation ranges (cut-off frequencies) is completely described. Conversely it is obvious that, by taking the quotient $z_1=(z_1z_2)^{1/2}:(z_2/z_1)^{1/2}$ and the product $z_2=(z_1z_2)^{1/2}\cdot(z_2/z_1)^{1/2}$, any pair of such functions $(z_2/z_1)^{1/2}$ and $(z_1z_2)^{1/2}$ as just described, gives physically realizable z_1 or $Z_1=Rz_1$, and z_2 or $Z_2=Rz_2$ of a filter with the prescribed transmitting and attenuation bands.

We may classify the filters into different "types" corresponding to the nature of their attenuation and transmission ranges. The four simplest and practically most important filter types are: low-pass, high-pass (both having one transmitting and one attenuation band), band pass (one transmitting and two attenuation bands), low-and-high-pass filter (two transmitting and one attenuation band). Both the first pair and the last pair are complemen-

tary (see above). Each type may be divided into an infinite number of classes which differ from each other by the number of zeros and poles of $(z_2/z_1)^{1/2}$ (or $(z_1z_2)^{1/2}$) in the transmitting bands (or attenuation bands). The following tables give the simplest classes for the low-pass filter type with cut-off frequence ω_1. The general law of building becomes quite clear from the table.

TABLE I. $(Z_2/Z_1)^{1/2}$ *for low pass filter.*

$$(1) \quad m \cdot \frac{\lambda}{(\lambda^2+\omega_1^2)^{1/2}}$$

$$(2) \quad m \cdot \frac{\lambda^2+\omega_a^2}{\lambda(\lambda^2+\omega_1^2)^{1/2}}$$

$$(3) \quad m \cdot \frac{\lambda(\lambda^2+\omega_a^2)}{(\lambda^2+\omega_b^2)(\lambda^2+\omega_1^2)^{1/2}}$$

$$(4) \quad m \cdot \frac{(\lambda^2+\omega_c^2)(\lambda^2+\omega_a^2)}{\lambda(\lambda^2+\omega_b^2)(\lambda^2+\omega_1^2)^{1/2}}$$

$$(5) \quad m \cdot \frac{\lambda(\lambda^2+\omega_c^2)(\lambda^2+\omega_a^2)}{(\lambda^2+\omega_d^2)(\lambda^2+\omega_b^2)(\lambda^2+\omega_1^2)^{1/2}}$$

$$(6) \quad m \cdot \frac{(\lambda^2+\omega_e^2)(\lambda^2+\omega_c^2)(\lambda^2+\omega_a^2)}{\lambda(\lambda^2+\omega_d^2)(\lambda^2+\omega_b^2)(\lambda^2+\omega_1^2)^{1/2}}$$

. . . . . .

. . . . . .

TABLE II. $(Z_1/Z_2)^{1/2}$ *for low pass filter.*

$$(\alpha) \quad \mu \cdot \frac{1}{(\lambda^2+\omega_1^2)^{1/2}}$$

$$(\beta) \quad \mu \cdot \frac{(\lambda^2+\omega_\alpha^2)}{(\lambda^2+\omega_1^2)^{1/2}}$$

$$(\gamma) \quad \mu \cdot \frac{\lambda^2+\omega_\alpha^2}{(\lambda^2+\omega_1^2)^{1/2}\,(\lambda^2+\omega_\beta^2)}$$

$$(\delta) \quad \mu \cdot \frac{(\lambda^2+\omega_\alpha^2)(\lambda^2+\omega_\gamma^2)}{(\lambda^2+\omega_1^2)^{1/2}\,(\lambda^2+\omega_\beta^2)}$$

$$(\epsilon) \quad \mu \cdot \frac{(\lambda^2+\omega_\alpha^2)(\lambda^2+\omega_\gamma^2)}{(\lambda^2+\omega_1^2)^{1/2}\,(\lambda^2+\omega_\beta^2)(\lambda^2+\omega_\delta^2)}$$

$$(\zeta) \quad \mu \cdot \frac{(\lambda^2+\omega_\alpha^2)(\lambda^2+\omega_\gamma^2)(\lambda^2+\omega_\epsilon^2)}{(\lambda^2+\omega_1^2)^{1/2}\,(\lambda^2+\omega_\beta^2)(\lambda^2+\omega_\delta^2)}$$

. . . . . .

. . . . . .

The parameters m, ω_a, ω_b, ω_c, $\cdots$, μ, ω_α, ω_β, ω_γ, $\cdots$ have to satisfy the conditions

$$m > 0, 0 < \cdots < \omega_c < \omega_b < \omega_a < \omega_1; \ \mu > 0, \ \omega_1 < \omega_\alpha < \omega_\beta < \omega_\gamma < \cdots \qquad (9)$$

in order that Z_1 and Z_2 may be physically realizable. The resonant and antiresonant frequencies ω_a, ω_b, $\cdots$, ω_α, ω_β, $\cdots$ in Tables I and II are written in their natural order of magnitude as indicated in (9).

The tables illustrate in the simplest type the above general remarks about the nature of the functions $(z_2/z_1)^{1/2}$ and $(z_1z_2)^{1/2}$. If we use the method above mentioned to derive reciprocal filters by replacing z_1 by $1/z_1$ and z_2

by $1/z_2$ or, what is the same, by replacing z_2/z_1 by z_1/z_2 and similarly z_1z_2 by $1/z_1z_2$ *the tables give all possible low-pass filter classes* (except for an interchange of Z_1 and Z_2 which is meaningless practically). The higher the class the greater is the number of the circuit elements (coils and condensers) which are necessary to realize the class. The example of the low-pass filter type illustrates quite clearly that each attenuation class (1, 2, 3, $\cdots$) may be combined with each image impedance class (α, β, γ, $\cdots$) (this is true for all types of filters). We use the class notation $n\alpha_m$ (where $\alpha_1 \equiv \alpha$, $\alpha_2 \equiv \beta$, $\alpha_3 \equiv \gamma$, $\cdots$) for a low-pass filter the attenuation class of which is n and the image impedance or characteristic impedance class of which is α_m.* For instance 3β means a low-pass filter of attenuation class 3 and of image impedance class β. By dividing and multiplying the corresponding terms from the table, there results, that for this particular class (and only for this class) the frequency characteristics Z_1 and Z_2 have the following expression:

$$Z_1 = \frac{\mu(\lambda^2 + \omega_b^2)(\lambda^2 + \omega_a^2)}{m\lambda(\lambda^2 + \omega_a^2)}$$
$$Z_2 = \frac{m\mu\lambda(\lambda^2 + \omega_a^2)(\lambda^2 + \omega_a^2)}{(\lambda^2 + \omega_b^2)(\lambda^2 + \omega_1^2)}. \tag{10}$$

In each class there are arbitrary parameters as m, ω_a, ω_b, $\cdots$, μ, ω_a, ω_β, $\cdots$ which may be arbitrarily chosen except for the restrictions made in (9). An important question in the designing of filters is the most economical determination of these parameters. The features of the attenuation and image impedance curves depend on the choice of these parameters. The form of characteristics which is in general most desirable for practical purposes is illustrated in Fig. 4 and 5.

Fig. 4 shows in curves 1, 2 and 3 frequency characteristics of the attenuation constant A_1 for the three lowest classes of a low-pass filter in the attenuation range. Abscissae are the normalized frequency $\Omega = \omega/\omega_1$, where ω_1 is the cut-off frequency of the low-pass filter. Ordinates are the attenuation constant

$$A_1 = \log_e \left| \frac{(z_2/z_1)^{1/2} + 1}{(z_2/z_1)^{1/2} - 1} \right|.$$

In the transmitting range not shown in Fig. 4, A_1 is zero. The dotted line $A_1 = 4.7$ indicates the practical requirement $A_1 \geqq A_{1\,\min} = 4.7$ for all frequencies $\Omega > k$, where k is the smallest value with this property. k decreases with increasing class-number ($k_3 < k_2 < k_1$ in Fig. 4). Or, if k is given, for $\Omega > k$ $A_{1\,\min}$ has the greatest possible value. We are interested practically only in a certain warranted minimum attenuation constant A_1 to the right of a given frequency, and do not want unnecessarily high attenuation with increasing ω. Therefore the described form of "Tschebyscheff approximation" of the

* A filter reciprocal to a filter of class $n\alpha_m$ in the sense defined above (see also appendix II) may be denoted by an additional star as $n\alpha_m{}^*$.

ideal requirement (infinite attenuation) in the attenuation range, is the most economical one. With this kind of approximation it is often possible to satisfy the same given attenuation requirements with a lower class than a known filter and to save coils and condensers in this way (such a comparison is made farther below).

Similarly, Fig. 5 shows in curves β and γ the frequency characteristics of the reduced image impedance $Z/R=(z_1z_2)^{1/2}$ in the transmitting band for two

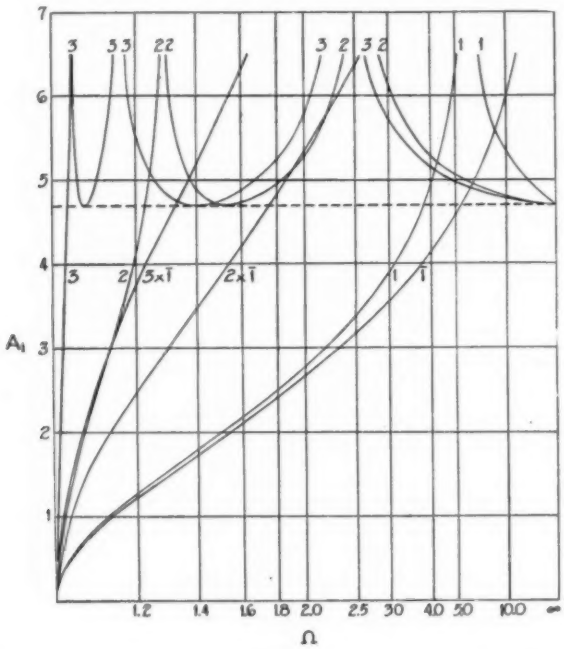

Fig. 4. Attenuation characteristics of a low pass filter.

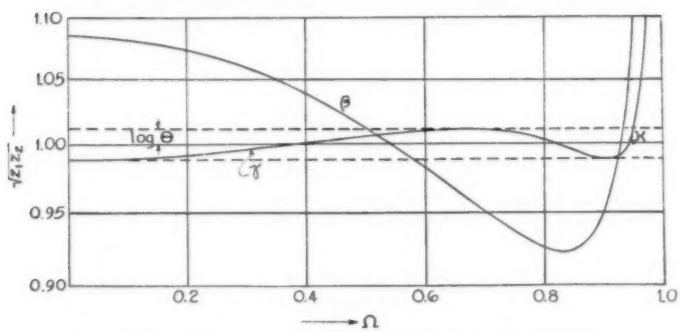

Fig. 5. Image impedance characteristics of a low pass filter.

of the lowest classes of a low-pass filter. Abscissae are again the normalized frequencies $\Omega=\omega/\omega_1$. The curves give again a "Tschebyscheff approximation" to the ideal requirement $(z_1z_2)^{1/2}=1$. In both curves β and γ, $\log_e (z_1z_2)^{1/2}$ has a minimum deviation from zero for $0\leqq\Omega\leqq\kappa$. The corresponding curve for the lowest image impedance class α which is not shown in Fig. 5 goes far below

the limits of the drawing. The minimum value $\log \Theta$ of $\log (z_1 z_2)^{1/2}$ decreases for a given κ with increasing class number but increases to infinity if κ goes to 1. Curve γ oscillates exactly between the limits $(z_1 z_2)^{1/2} = \Theta$ and $(z_1 z_2)^{1/2} = 1/\Theta$ (dotted lines) in the approximation interval. This is the typical form of a characteristic of "Tschebyscheff approximation." The advantages are of the same kind as in the case of attenuation characteristics.

k or κ may be given as near as wanted to the normalized cut-off frequency $\Omega = 1$, but not equal to 1. Then by these k or κ the values $A_{1\,\min}$ or Θ are determined uniquely, as also are the "Tschebyscheff parameters" m, ω_a, ω_b, $\cdots$, (μ/R), ω_α, ω_β, $\cdots$, for each class, where $A_{1\,\min} \to \infty$ and $\Theta \to 1$ with increasing class number.

Indeed, the most important fact is that the physical realizability (9) is satisfied by the Tschebyscheff parameters.*

Analogous considerations are valuable in the cases of high-pass, band-pass and low-and-high pass filters which are given in detail in the papers cited above.

There is another way to solve the problem of satisfying approximately the ideal requirements (6) and (7) by functions of the kind (1) by means of a Poisson integral representation. This method gives a very clear proof of the possibility of the solution of our problem, showing the limiting case where the class-number goes to infinity and the approximation becomes ideal. This method of approximation may be of practical value in the cases of higher filter types (for instance more than two bands of free transmission) where the explicit determination of the parameters for Tschebyscheff approximation involves considerable difficulty. We take again the simple example of the low-pass filter type with cut-off frequency ω_1.

A Poisson integral representation of $\lambda(\lambda^2+\omega_1^2)^{1/2}$ gives us the identity

$$\lambda(\lambda^2 + \omega_1^2)^{1/2} \cdot \frac{1}{\pi} \int_0^{\omega_1^2} \frac{dx}{(x(\omega_1^2 - x))^{1/2}(\lambda^2 + x)} = 1. \tag{11}$$

It is essential that the factor of $1/\lambda^2+x$ in the integrand be always positive. After cutting the imaginary axis along $|\lambda| \leqq \omega_1$ we are only interested in that branch of $(\lambda^2+\omega_1^2)^{1/2}$ in the right λ half-plane which is positive for $\lambda > 0$. By approximating the integral by a sum we get an approximation of 1 uniform in the whole λ-plane after excluding a domain which contains the points $|\lambda| \leqq \omega_1$ of the imaginary λ-axis in its interior. But this sum has the form

$$S = \sum_{\nu=1}^{n} \frac{a_\nu}{\lambda^2 + b_\nu}$$

where $a_\nu > 0$, $b_\nu \geqq 0$. It follows at once that S can be written in the form

* The mathematical treatment for low-pass, high-pass, band-pass and low-and-high-pass filter is given in a forthcoming paper "Ein Interpolationsproblem mit Funktionen mit positivem Realteil". A detailed numerical elaboration with consideration of the resistances for all named types of filters was given in a book "Siebschaltungen" (a collection of tables and curve sheets, V.D.I.-Verlag 1931) for the purposes of practical design.

$$m\frac{(\lambda^2+\omega_t^2)\cdots(\lambda^2+\omega_b^2)}{(\lambda^2+\omega_u^2)(\lambda^2+\omega_s^2)\cdots(\lambda^2+\omega_a^2)}$$

where m and the ω's satisfy the condition (9). Because it is possible to get approximations of the integral by taking $\omega_u=0$ or $\omega_a=\omega_1$ or both, all possible $(z_2/z_1)^{1/2}$ classes as given in Table I, including the reciprocal classes, may be obtained on the left side of (11) by replacing the integral there by a sum S.

Starting with the analogous identity

$$(\lambda^2+\omega_1^2)^{1/2}\cdot\frac{1}{\pi}\int_{\omega_1^2}^{\infty}\frac{dx}{(x-\omega_1^2)^{1/2}(\lambda^2+x)}=1 \tag{12}$$

corresponding approximations of the ideal requirement (6) may be again obtained by approximating the integral by a sum. Any class of all possible $(z_1z_2)^{1/2}$ classes is given in Table II (there $(Z_1Z_2)^{1/2}=R\cdot(z_1z_2)^{1/2}$), including the reciprocal classes, arises in this approximation method.

For the practical design of a filter the following circumstance is useful: The total or working attenuation A as given in formula (4) is nearly the superposition of a term A_1 (the attenuation constant) which depends only on $(z_2/z_1)^{1/2}$, and a second term A_2 which depends only on $(z_1z_2)^{1/2}$. As a correction only, it is sometimes necessary to consider a remaining term A_3, which depends on both $(z_2/z_1)^{1/2}$ and $(z_1z_2)^{1/2}$. In formulas:

$$\begin{aligned} A &= A_1+A_2+A_3 \\ A_1+iB_1 &= \log_e\frac{\left(\frac{z_2}{z_1}\right)^{1/2}+1}{\left(\frac{z_2}{z_1}\right)^{1/2}-1} \\ A_2 &= \log_e\left|\frac{((z_1z_2)^{1/2}+1)^2}{4(z_1z_2)^{1/2}}\right| \\ A_3 &= \log_e\left|1-e^{-2[A_1+A_1'+i(B_1+B_1')]}\right| \end{aligned} \tag{13}$$

where

$$A_1'+iB_1'=\log_e\frac{(z_1z_2)^{1/2}+1}{(z_1z_2)^{1/2}-1}. \tag{14}$$

It may be easily verified that these formulas give the result (4). This kind of separation corresponds to that given by O. J. Zobel in his paper about a chart method of designing wave filters, B.S.T.J., 1924. It is superfluous to go into the detail of the practical design of filters here since I have already published a paper for this purpose (see above).

Now after describing the principles of the theory and design of general symmetrical wave filters it remains to show more clearly by a series of examples that all known symmetrical filters are only very special cases of filters and scarcely the most economical ones. First we shall derive some rules that

are useful in determining the characteristics Z_1 and Z_2 and the class of a given symmetrical four-terminal network.

Though there is no particular difficulty in calculating Z_1 and Z_2 by determining the short-circuit and open-circuit impedance Z_s and Z_0 of the whole network using the formula

$$\left.\begin{matrix} Z_1 \\ Z_2 \end{matrix}\right\} = Z_0 \pm [Z_0(Z_0 - Z_s)]^{1/2} \tag{14a}$$

this calculation process may be replaced in most filters used heretofore by calculating the characteristics of very simple parts of the whole filter. This is because it usually shows some chain or symmetry structure.[11]

The series combination of two matched (see page 243) symmetrical four-terminal networks as 1, 1′; 2, 2′ with characteristics Z_1, Z_2 and 2, 2′; 3, 3′ with characteristics Z_3, Z_4 in Fig. 6a is again a symmetrical four-terminal network. The condition of matching is $Z_1Z_2 = Z_3Z_4$. Because the total propagation constant Γ_3 is the sum of the propagation constants Γ_1 and Γ_2 of both parts as explained in the beginning, we get, by adding the expressions (2) of both parts for the characteristics Z_5 and Z_6 of the resulting network (Fig. 6b), the expression:

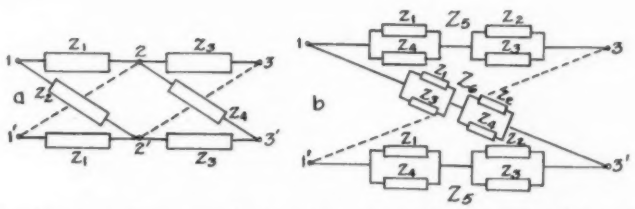

Fig. 6. Equivalence of two matched lattice type sections in series (a), and a single lattice type section (b).

$$\left(\frac{Z_6}{Z_5}\right)^{1/2} = \frac{1 + (Z_2/Z_1)^{1/2}(Z_4/Z_3)^{1/2}}{(Z_2/Z_1)^{1/2} + (Z_4/Z_3)^{1/2}}. \tag{15}$$

As mentioned in the beginning the image impedance of the whole network is equal to the image impedances on the free ends of the connected sections:

$$(Z_5Z_6)^{1/2} = (Z_1Z_2)^{1/2} = (Z_3Z_4)^{1/2}. \tag{16}$$

Z_5 may be built by a parallel combination of Z_1 and Z_4 in series with a parallel combination of Z_2 and Z_3, and Z_6 is built by a parallel combination of Z_1 and Z_3 in series with a parallel combination of Z_2 and Z_4. (Fig. 6b.) Nevertheless this theorem is more of theoretical interest than of practical use for our purpose, the structure Fig. 6b having also superfluous elements.

It is possible to show from (15) that by series combination of matched symmetrical filter sections the attenuation class numbers are added, which

[11] A general physical construction method of Z_1, Z_2 for an arbitrary symmetrical four-terminal network was given in Fig. 7a of my paper "Vierpole," E.N.T 1929, p. 279. But the other methods given here abridge the calculation of Z_1 and Z_2 where they are applicable.

means that the attenuation class number of the resulting network is the sum of the attenuation class numbers of the parts.[12]

The formula (15) considers the orientation of the resulting four-terminal network (Fig. 6b) too. This means that if 1 in Fig. 6b corresponds to 1 in Fig. 6a and 1′ to 1′, then, if (15) and (16) are true, 3 corresponds to 3 and 3′ to 3′ (not 3 to 3′ or 3′ to 3). A reversal of the orientation of one of the two sections (for instance permutation of Z_1 and Z_2) of Fig. 6a changes the orientation of the resulting network (Fig. 6b) as well. The additive property of the attenuation class number, if matched symmetrical filter sections are put in series, is right too, including the orientation. But it is obvious now that if the position of $Z_1(Z_3, Z_5)$ in the lattice section is fixed as indicated in Fig. 6 and we want to keep the additive property just mentioned, including the orientation, the heading of Table I is not arbitrarily Z_1/Z_2 or Z_2/Z_1, but must be chosen Z_2/Z_1.

A first example of the foregoing is the chain Fig. 7b, the three sections of which are equal. They are not merely symmetrical four-terminal networks

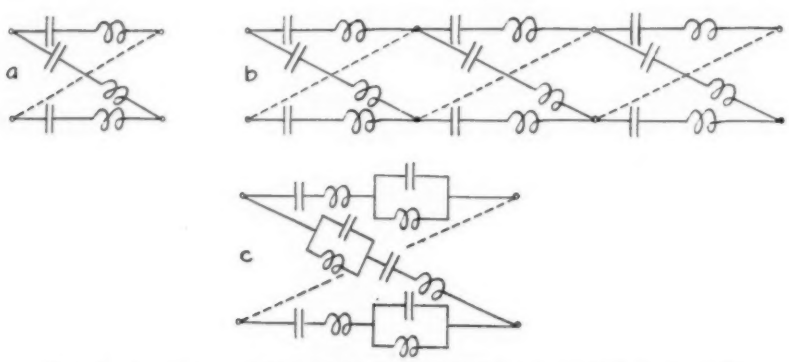

Fig. 7. (a) Band pass filter of class $1b^*$. (b) Band pass filter of class $3b^*$ having three identical sections a.

but are directly given as lattice structures. The single section (Fig. 7a) represents a band-pass filter which transmits in the interval

$$\omega_{-1} < \omega < \omega_1,$$

with

$$\omega_{-1} = \frac{1}{(L_1C_1)^{1/2}}$$

and

$$\omega_1 = \frac{1}{(L_2C_2)^{1/2}},$$

if we assume $L_1C_1 > L_2C_2$. The single section belongs in my notation (see the paper "Siebschaltungen," V. D. I. Verlag) to class 1b* with attenuation class 1*:

[12] This is true not only for low-pass filters (Table I), but also for high-pass and band-pass filters (not low-and-high-pass filters) if the notation of my book "Siebschaltungen" is used.

$$\left(\frac{Z_1}{Z_2}\right)^{1/2} = \left(\frac{L_1}{L_2}\right)^{1/2} \cdot \frac{(\lambda^2 + \omega_{-1}^2)^{1/2}}{(\lambda^2 + \omega_1^2)^{1/2}} \tag{17}$$

and image impedance class b^*:

$$\frac{1}{(Z_1Z_2)^{1/2}} = \frac{1}{(L_1L_2)^{1/2}} \cdot \frac{\lambda}{((\lambda^2 + \omega_{-1}^2)(\lambda^2 + \omega_1^2))^{1/2}}. \tag{18}$$

It follows from the rule given above that the filter chain Fig. 7b with three equal sections of class 1b* represents a filter of the class 3b*, which is equivalent to the single lattice type section Fig. 7c, the total A_1 characteristic being determined by

$$\left(\frac{Z_1}{Z_2}\right)^{1/2} = m \frac{(\lambda^2 + \omega_1^2)^{1/2}(\lambda^2 + \omega_a^2)}{(\lambda^2 + \omega_{-a}^2)(\lambda^2 + \omega_1^2)^{1/2}} \tag{19}$$

as may be easily derived by help of the formula (15). The image impedance is again (18).[13]

If a symmetrical four-terminal network is not only electrically, but also structurally, symmetrical, and consists of two not necessarily symmetrical four-terminal networks that are not related by mutual inductance, as illustrated in Fig. 8, Z_1 is the impedance of the two-terminal network which is obtained by short circuiting one half in the middle. Simularly Z_2 is obtained as open circuit impedance of the one half (measured on 1, 1′ or 2, 2′).[14] A generalization of this "symmetry theorem" states that the analogous construction of Z_1 and Z_2 remains true if the number of connecting points between both parts, one being the mirror image of the other, is arbitrary.[15] This theorem, too, is very valuable in determining the class of a given filter.

By this method we find at once for the well-known filter section Fig. 9a

$$Z_1 = \frac{L\lambda}{LC\lambda^2 + 1}$$

$$Z_2 = 1/C\lambda.$$

The filter section Fig. 9b is obtained from Fig. 9a by changing the orientation (changing the terminals 2 and 2′) or interchanging Z_1 and Z_2. By calculating Z_2/Z_1 and Z_1Z_2 it is found that these two simple filters belong to the low-pass

[13] The filter Fig. 7b was recently in use for carrier telegraphy purposes, opposite coils being connected as two windings on one core. I have shown (see "Siebschaltungen") that by designing another filter of the same class similar in structure to Fig. 7c, it is possible to save 33 percent of the coils and 50 percent of the condensers, at the same time essentially improving the attenuation characteristic. The reason for this last circumstance is the avoiding of unnecessary restrictions in the choice of the arbitrary parameters (m, (m, ω_{-a}, ω_a) caused by the chain structure. This becomes clearer by the next example further below for a low-pass filter.

[14] Proved in my paper in Archiv für Elektrotechnik Dec. 1926.

[15] A. C. Bartlett, Phil. Mag., 902–907, Nov. 1927, "An Extension of a Property of Artificial Lines;" see also the book of A. C. Bartlett "The Theory of Electrical Artificial Lines and Fillers", p. 28.

filter class 1α, but only the filter Fig. 9b is directly written down in the Tables I and II without interchanging Z_1 and Z_2. The connection between the arbitrary parameters appearing in the tables and the circuit elements is in this case

$$m = 1 \quad \mu = 1/C \quad \omega_1^2 = 1/LC. \tag{20}$$

According to the additive property of the attenuation class number in filter chains a chain of s equal sections of the kind Fig. 9a or 9b is again a low-pass filter which belongs to the attenuation class s while the image impedance class is the same as of the single section. Taking sections of the kind Fig. 9b,

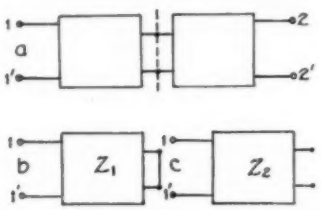

Fig. 8. (a) Structurally symmetrical four-terminal network. (b), (c) Z_1 and Z_2 of the equivalent canonical symmetrical four-terminal network of Fig. 3.

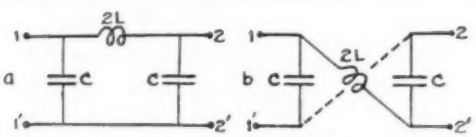

Fig. 9. Low pass filter of class 1α.

the repeated application of (15) gives for a chain with 1 to 4 sections the following $(Z_2/Z_1)^{1/2}$ classes:

$$\begin{aligned}
&(1) \quad \frac{\lambda}{(\lambda^2+\omega_1^2)^{1/2}} \\
&(2) \quad \frac{\lambda^2+\frac{\omega_1^2}{2}}{\lambda(\lambda^2+\omega_1^2)^{1/2}} \\
&(3) \quad \frac{\lambda(\lambda^2+\frac{3}{4}\omega_1^2)}{\left(\lambda^2+\frac{\omega_1^2}{4}\right)(\lambda^2+\omega_1^2)^{1/2}} \\
&(4) \quad \frac{\left(\lambda^2+\frac{(2-(2)^{1/2})\omega_1^2}{4}\right)\left(\lambda^2+\frac{(2+(2)^{1/2})\omega_1^2}{4}\right)}{\lambda\left(\lambda^2+\frac{\omega_1^2}{2}\right)(\lambda^2+\omega_1^2)^{1/2}}.
\end{aligned} \tag{21}$$

The attenuation constant A_1 for (2) must be twice that for (1), that for (3) must be 3 times that for (1) and so on. Since the function (1) is 1 only for

$\lambda = i\omega \to \infty$ the $\Omega = \omega/\omega_1 \div A_1$ – characteristic for such a chain with an arbitrary number of sections has always one point of infinity for $\Omega = \infty$ and no other point of infinity. The attenuation curves of a chain with 1 to 3 sections are drawn in Fig. 4 and denoted by $\bar{1}$, $2\times\bar{1}$, $3\times\bar{1}$. The attenuation curve 2 of class 2, which gives a Tschebyscheff approximation, has a smaller normalized frequency $\Omega = k_2$ above which the minimum attenuation $A_1 = 4.7$ is warranted, than curve $3\times\bar{1}$ of the higher class 3. This makes very obvious the advantage of the use of Tschebyscheff parameters.

It is of some interest to see that the same $(Z_2/Z_1)^{1/2}$ functions corresponding to the chain just treated may be obtained by quite another approximation method. We consider the periodic continued fraction

$$(\lambda^2 + \omega_1^2)^{1/2}/\lambda = 1 + \frac{1}{|2\lambda^2/\omega_1^2} + \frac{1}{|2} + \frac{1}{|2\lambda^2/\omega_1^2} + \frac{1}{|2} + \cdots \tag{22}$$

and take any approximating fraction (cutting off the infinite remainder of the whole fraction after a certain point) multiplied by $\lambda/(\lambda^2+\omega_1^2)^{1/2}$. Because $(\lambda^2+\omega_1^2)^{1/2}/\lambda - 1$ is a Stieltjes continued fraction[16] as function of λ^2 it is known in advance that the functions approximating 1 obtained in this way are not only formally of the form given in Table I, but also satisfy the conditions (9). Now the functions found in this way happen to be just the same as the

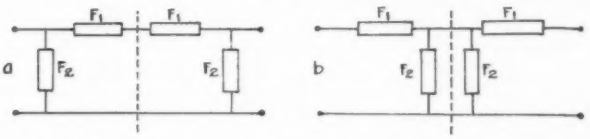

Fig. 10. (a) Symmetrical π-network. (b) Symmetrical T-network.

$(Z_2/Z_1)^{1/2}$ functions corresponding to the chain treated before. This may be easily verified for the first 4 classes given in (21). This gives a good illustration of the fact that the chain method is only one of several methods of approximating the ideal requirement (7).[17]

In the following, the influence of the terminations or end circuits which are in use for changing and improving the image impedance characteristic, on the class of the wave filter may be discussed. For the general π-structure network, Fig. 10a, where F_1 and F_2 signify general impedances, the following relations hold according to the symmetry theorem:

$$\left(\frac{Z_2}{Z_1}\right)^{1/2} = \left(\frac{F_1 + F_2}{F_1}\right)^{1/2}$$
$$\frac{1}{(Z_1 Z_2)^{1/2}} = \left(\frac{1}{F_2}\left(\frac{1}{F_1} + \frac{1}{F_2}\right)\right)^{1/2}. \tag{23}$$

[16] The Stieltjes continued fractions play an important role in the theory of circuits. T. C. Fry, Bulletin of the American Mathematical Society, 1929, and W. Cauer, Jahresbericht der Deutschen Mathematikervereinigung, 1929, in which papers further literature may be found.

[17] Indeed my first intention in studying continued fractions like (22) was to find a new form of approximation of the ideal requirement (7), and I found later on the coincidence in the results with the chain method.

If both halves of Fig. 10a are combined in the other way, the network Fig. 10b arises. For this dual network we find analogously

$$\left(\frac{Z_2}{Z_1}\right)^{1/2} = \left(\frac{F_1 + F_2}{F_1}\right)^{1/2}$$
$$(Z_1 Z_2)^{1/2} = (F_1(F_1 + F_2))^{1/2}. \tag{24}$$

That it has the same attenuation characteristic $(Z_2/Z_1)^{1/2}$ as the foregoing network, was obvious in advance because the total attenuation constant, according to the definition given in the beginning of this paper, is in both cases Fig. 10a and Fig. 10b the sum of the Γ's for both halves. If $F_2 = R^2/F_1$, where R is a positive constant as before, as is the case in the "constant k" wave filter of Zobel ($k = R$), the $(Z_1 Z_2)^{1/2}$ for the T-network Fig. 10b is equal to $R^2/(Z_1 Z_2)^{1/2}$ for the π-network Fig. 10a. In the case where F_2 is the reciprocal

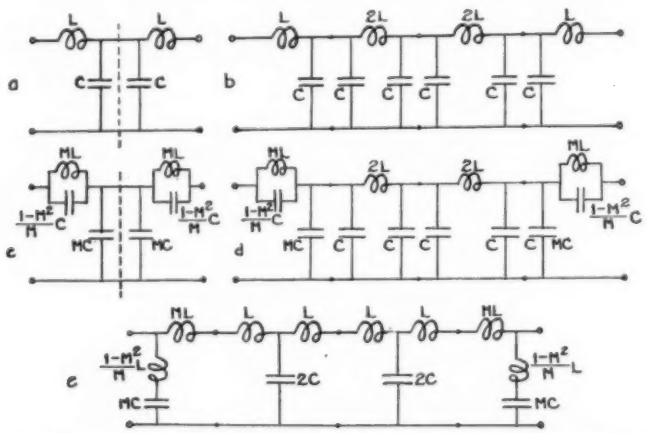

Fig. 11. (a) Low pass filter of class $1\alpha^*$. (b) Low pass filter of class $3\alpha^*$. (c) Low pass filter of class $1\beta^*$. (d) Low pass filter of class $3\beta^*$. (e) Low pass filter of class 3β.

of F_1 (except for the positive constant R^2), it follows that the networks Fig. 10a and Fig. 10b are reciprocal without regard to the orientation (compare appendix II). More especially, if Fig. 10a is a filter such as that of Fig. 9a ($F_1 = L\lambda$, $F_2 = 1/C\lambda$, $R^2 = L/C$ low-pass filter class 1α), Fig. 10b is a reciprocal filter and belongs to the reciprocal class $1\alpha^*$ (Fig. 11a). Our construction has introduced a new image impedance characteristic. The halves of Fig. 11a may be taken as terminals of a chain with several sections like Fig. 9a, the condition of matching being satisfied in this case. As an example, Fig. 11b results and is of class $3\alpha^*$. The attenuation characteristic $(Z_1/Z_2)^{1/2}$ of this filter is given by (21) (3), $\omega_1^2 = 1/LC$, and the image impedance characteristic $1/(Z_1 Z_2)^{1/2}$ by Table II, α, $\omega_1^2 = 1/LC$, $\mu = 1/L$. Of course the filter Fig. 11b may also be considered as a series combination of three equal sections Fig. 11a.

Now, as a typical example of terminations of a filter, Zobel's "M-type" may be considered. A symmetrical π-network (Fig. 10a) may be transformed with a constant positive parameter M in the more complicated π-net-

work Fig. 12a which is physically realizable if $M<1$. The meaning of this transformation becomes very clear by introducing our Z_1 and Z_2 for Fig. 10a and the corresponding Z_1^* and Z_2^* for Fig. 12a. The symmetry theorem shows at once that

$$Z_1^* = MZ_1$$

and

$$Z_2^* = M^{-1}Z_2.$$

That means, the image impedance $(Z_1Z_2)^{1/2}$ is the same for both networks and

$$\left(\frac{Z_2^*}{Z_1^*}\right)^{1/2} = M^{-1}\left(\frac{Z_2}{Z_1}\right)^{1/2} \tag{25}$$

(compare (23)): It follows in the case of a filter that the attenuation class is also unchanged. If we put together both halves of Fig. 12a in the other way as shown for the case $F_1=L\lambda$, $F_2=1/C\lambda$ in Fig. 11c, the attenuation charac-

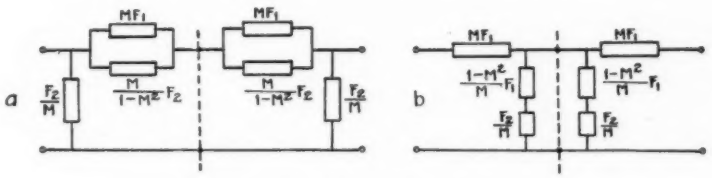

Fig. 12. M-type structures.

teristic is not changed, as explained in the case of Fig. 10a and 10b. But according to the symmetry theorem, the image impedance becomes now

$$(Z_1'Z_2')^{1/2} = \frac{F_2(F_1(F_1+F_2))^{1/2}}{(1-M^2)F_1+F_2} \tag{26}$$

and more especially for Fig. 11c

$$(Z_1'Z_2')^{1/2} = \frac{(\lambda^2+\omega_1^2)^{1/2}}{C(1-M^2)\left(\lambda^2+\dfrac{\omega_1^2}{1-M^2}\right)}, \quad \omega_1^2 = \frac{1}{LC}. \tag{27}$$

Taking the right half of Fig. 12a, this network has on the left side an image impedance (26) and on the right side an image impedance (23). This half may be the left hand "M-type" terminal of a chain with matched sections, the other half the right hand terminal. If we take the special case $F_1=L\lambda$, $F_2=1/C\lambda$, the low-pass filter of Fig. 11d with two sections like Fig. 9a is obtained. Fig. 11c belongs to class $1\beta^*$, according to (25), (23), (27) and Tables I and II, where

$$\omega_1^2 = \frac{1}{LC}, \quad m = M, \quad \mu = C(1-M^2), \quad \omega_\alpha^2 = \frac{\omega_1^2}{1-M^2}. \tag{28}$$

Filter Fig. 11d having the same image impedance as Fig. 11c belongs to the class $3\beta^*$, according to the additive property of the attenuation class in chains with matched sections.[18]

The one disadvantage of such terminating circuits which improve the image impedance characteristic is that their use does not allow one to choose independently the $(Z_2/Z_1)^{1/2}$- and $(Z_1Z_2)^{1/2}$-characteristic. For instance, the M-type terminals of Fig. 11d give (see (28)) the partial attenuation term

$$\left(\frac{Z_1}{Z_2}\right)^{1/2} = \frac{M\lambda}{(\lambda^2 + \omega_1^2)^{1/2}},$$

which becomes 1 for

$$\Omega = \frac{\omega}{\omega_1} = \frac{1}{(1 - M^2)^{1/2}} \tag{29}$$

or $\Omega = 1.25$ for the practical value $M = 0.6$. It follows that with Zobels M-type, $M = 0.6$, a normed frequency $\Omega = 1.25$ of infinite attenuation constant A_1 is connected. Because (27) becomes infinite at the value (29), the working attenuation A becomes infinite there whether A_1 is infinite at the value (29) or not, according to (4) or the third formula (13).

Quite similar considerations relate to the dual "M-type". The network Fig. 12b corresponds to the network Fig. 10b, as Fig. 12a corresponds to Fig. 10a. Again the relations

$$Z_1^* = MZ_1$$

and

$$Z_2^* = M^{-1}Z_2$$

are true if Z_1, Z_2 denote the characteristics of Fig. 10b and Z_1^*, Z_2^* those of Fig. 12b. If these networks are filters, both have the same attenuation class and $(Z_2^*/Z_1^*)^{1/2}$ for Fig. 12b agrees with the corresponding expression for Fig. 12a. The left half of Fig. 12b has the left hand image impedance

$$W_1 = (F_1(F_1 + F_2))^{1/2} \tag{30}$$

and the right hand image impedance

$$W_2 = [(1 - M^2)F_1 + F_2]\left(\frac{F_1}{F_1 + F_2}\right)^{1/2}. \tag{31}$$

The last expression is equal to $R^2/(Z_1'Z_2')^{1/2}$ where the value $(Z_1'Z_2')^{1/2}$ is taken from (26), if

$$F_2 = \frac{R^2}{F_1}$$

[18] For the calculation of the $(Z_2/Z_1)^{1/2}$-characteristic of the whole chain Fig. 11d the left hand terminal may be omitted and replaced by a congruent added section on the right side of the filter. Then the repeated application of (15) gives the total $(Z_2/Z_1)^{1/2}$.

holds. It follows, as an example, that the filter network Fig. 11e belongs to the class 3β, since Fig. 11d belongs to class $3\beta^*$.

Finally the class and the characteristics Z_1 and Z_2 of the more complicated filter Fig. 13 may be determined. This was calculated by O. J. Zobel in his paper in the B.S.T.J. 1923, p. 30, and is mentioned also as the chief example in the book of T. E. Shea, "Transmission Networks and Wave Filters", p. 323. This network consists of two matched symmetrical four-terminal networks in the middle, called X_1 and XI_1 in the paper of Zobel's just mentioned, in series with which two matched half sections IV_1 are connected, one at each end. These terminals are an equivalent form of the M-type explained before in Fig. 12b. There is no difficulty in determining the class and the $(Z_2/Z_1)^{1/2}$- and $(Z_1Z_2)^{1/2}$-characteristic using the symmetry theorem and the rule (15). It follows that the network belongs to the band-pass class $6d$ in the notation of my paper "Siebschaltungen". The formulas are derived in detail in appendix III.

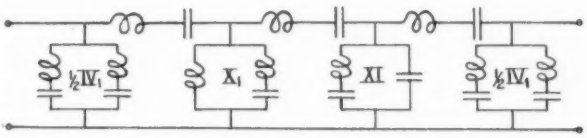

Fig. 13. O. J. Zobel's example of a band pass filter of class $6d$.

In the paper "Siebschaltungen" it is shown that by using Tschebyscheff parameters the same practical task which is solved by the network Fig. 13 of class $6d$ can be solved by a non-equivalent network of class $5d$ or $5d^*$, that is by a lower class and therefore by fewer coils and capacities.

The aim of this paper was to give an introduction into the new theory and the new design methods by illustrating the connection with the known filters; it does not attempt to give a complete statement of the theory or of the question of practical design. Reference is made once more to the papers cited above.

Appendix I. Total Attenuation

The total attenuation of a four-terminal network F which connects a sending apparatus S and a receiving apparatus E both of equal constant resistances R is defined as

$$A = \log_e \left| J'/J_2 \right| ,$$

J' denoting the current which flows in the receiving apparatus if S and E are immediately connected, J_2 being the corresponding current if the four-terminal network F is put between S and E. If V is the electromotive force in S we have

$$J' = V/2R.$$

If F is symmetrical—this is the only case we are interested in here—J_2 may be expressed by V, R and the characteristics Z_1 and Z_2 (compare Fig. 14),

using for instance the following theorem of Kirchhoff: "The current J_2 produced in branch 2 by a voltage V in branch 1 of a network with n independent meshes without mutual inductances is equal to $V\Delta_1/\Delta$, where Δ_1 is the algebraic sum of the impedance products of all systems of $(n-1)$ branches, so that after their removal one mesh only remains going through branch 1 and 2, the term in the sum being taken with plus sign if the sense of V and J_2 is the same in the remaining mesh, otherwise with minus sign, and where Δ is the sum of the impedance products of all systems of n branches such that after their removal no mesh remains."

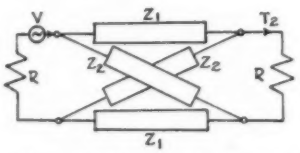

Fig. 14. Network illustrating the calculation of the total (working) attenuation of an arbitrary symmetrical four-terminal network.

In our case $n=3$, and we get at once (Fig. 14)

$$\Delta_1 = Z_2^2 - Z_1^2$$

$$\Delta = 2(Z_1 + Z_2)Z_1Z_2 + 2(Z_1^2 + Z_2^2)R + 4Z_1Z_2R + 2(Z_1 + Z_2)R^2.$$

It follows that

$$A = \log_e \left|\frac{J'}{J_2}\right| = \log_e \left|\frac{V}{2R}\cdot\frac{\Delta}{V\Delta_1}\right| = \log_e \left|\frac{\left(1+\frac{Z_1}{R}\right)\left(1+\frac{Z_2}{R}\right)}{\frac{Z_2}{R}-\frac{Z_1}{R}}\right|$$

which is indeed the formula (4).

Appendix II. Reciprocal Networks

For a $2q$-terminal network the following equations are valuable and characteristic if electromotive forces $V_1, V_2, \cdots, V_q$ are put between the pairs of terminals which have the corresponding subscript and $J_1, J_2, \cdots, J_q$ denote the currents through these pairs of terminals:

$$J_1 = \alpha_{11}V_1 + \alpha_{12}V_2 + \cdots + \alpha_{1q}V_q$$

$$J_2 = \alpha_{21}V_1 + \alpha_{22}V_2 + \cdots + \alpha_{2q}V_q$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$J_q = \alpha_{q1}V_1 + \alpha_{q2}V_2 + \cdots + \alpha_{qq}V_q.$$

The matrix of the coefficients (α_{rs}) is symmetrical $(\alpha_{rs}=\alpha_{sr})$. The elements are functions of the frequency parameter $\lambda=i\omega$. The necessary and sufficient conditions for the α_{rs} to correspond to a physically realizable passive network are not discussed here. Only the following theorem may be stated:

If (α_{rs}) represents the characteristic matrix of a physically realizable 2q-terminal network, and the determinant $|\alpha_{rs}|$ is not identically zero, then the reciprocal matrix $(\alpha_{rs})^{-1}$ always belongs to a physical realizable 2q-terminal network too[19] *(and also, instead of the reciprocal matrix such a matrix multiplied with an arbitrary positive constant R^{-2}).* We may agree to define a $2q$-terminal network with characteristic matrix (α_{st}) as reciprocal to another $2q$-terminal network with matrix (β_{st}) if (β_{st}) equals $(\alpha_{st})^{-1}$ except a positive constant factor R^{-2}.

In the special case of a four-terminal network the following equations hold:

$$J_1 = \alpha_{11}V_1 + \alpha_{12}V_2$$
$$J_2 = \alpha_{21}V_1 + \alpha_{22}V_2,$$

where $\alpha_{12}=\alpha_{21}$. Then the corresponding equations are true for the *reciprocal four-terminal network* if α is written for $\alpha_{11}\alpha_{22}-\alpha_{12}^2$:

$$J_1 = \frac{R^{-2}\alpha_{22}}{\alpha} V_1 - \frac{R^{-2}\alpha_{21}}{\alpha} V_2$$
$$J_2 = -\frac{R^{-2}\alpha_{12}}{\alpha} V_1 + \frac{R^{-2}\alpha_{11}}{\alpha} V_2.$$

If the further specialization is made to a symmetrical four-terminal network where $\alpha_{11}=\alpha_{22}$, the α's are expressed by the characteristics Z_1 and Z_2 of the filter network by means of

$$Z_1 = (\alpha_{11} + \alpha_{12})^{-1}$$
$$Z_2 = (\alpha_{11} - \alpha_{12})^{-1}.$$

The corresponding characteristics Z_1^* and Z_2^* for the reciprocal network are similarly

$$Z_1^* = \left(\frac{R^{-2}\alpha_{11}}{\alpha} - \frac{R^{-2}\alpha_{12}}{\alpha}\right)^{-1} = \frac{R^2\alpha}{\alpha_{11} - \alpha_{12}} = \frac{R^2}{Z_1}$$
$$Z_2^* = \left(\frac{R^{-2}\alpha_{11}}{\alpha} + \frac{R^{-2}\alpha_{12}}{\alpha}\right)^{-1} = \frac{R^2\alpha}{\alpha_{11} + \alpha_{22}} = \frac{R^2}{Z_2}.$$

These formulas show that the definition given in the text for reciprocal symmetrical filters is exactly a special case of the more general kind of reciprocity defined in the appendix. In order that this may be true for instance for Z_1 and Z_2 given by Table I and II, it is necessary indeed to replace not only $(Z_1Z_2)^{1/2}$ by $R^2/(Z_1^*Z_2^*)^{1/2}$, but also to replace $(Z_2/Z_1)^{1/2}$ by $(Z_1^*/Z_2^*)^{1/2}$ to obtain a reciprocal filter network. Without making the latter step Z_1^* and Z_2^* would be interchanged and the new network obtained would have another orientation than the reciprocal network. If we do not care about the orientation the equivalence of the following definition to the previous one of reciprocal four-terminal networks may be easily verified (compare "Vierpole,"

[19] This theorem is proved in my paper in the Mathematische Annalen, Bd. 105 (1931).

E.N.T. 1929, p. 281): "Two four-terminal networks are reciprocal if the open circuit characteristics of the first network are reciprocal (except for a positive constant factor R^2) to the corresponding short circuit characteristics of the second network and vice versa." If any three of these conditions are satisfied the fourth is satisfied as a consequence. In the case of a symmetrical four-terminal network there is only one open circuit impedance Z_0 and one short circuit impedance Z_s. Here indeed we get at once from (14α) by writing $Z_0^* = R^2/Z_s$ and $Z_s^* = R^2/Z_0$

$$\left.\begin{matrix} Z_1^* \\ Z_2^* \end{matrix}\right\} = \frac{R^2}{Z_0 \pm (Z_0(Z_0 - Z_s))^{1/2}}.$$

Appendix III. Class and Z_1, Z_2-Characteristics of the Filter Network of Fig. 13.

The parameters of the band-pass filter elements (Fig. 13) of Zobel are given in the Bell System Technical Journal, 1923, page 30. Because the circuit elements are not given directly, reference must be made to the paper cited.

A slight analytical simplification is possible for the present purpose by replacing the two sections X_1 and XI_1 by IV_1 and VII_1 according to Zobel's formula (44e). (The original structure may, however, be more desirable from the standpoint of magnitudes of the elements.) What we have to do now is to combine the $(Z_2/Z_1)^{1/2}$ characteristics of the three symmetrical sections IV_1, $IV_{1,M}$, VII (using formula (15)) and to calculate the $(Z_1Z_2)^{1/2}$ for the section $IV_{1,M}$ which is the same as the external image impedances of the whole chain. The critical frequencies, in radians per second, and R are (p. 30, l.c.)

$$\omega_{-1} = 2\pi \cdot 4000$$
$$\omega_1 = 2\pi \cdot 7000$$

$$R = 600 \text{ ohms}.$$

The parameters of the IV_1 section are given by (p. 30 and 41, l.c.)

$$m_1 = \frac{\omega_{-1}\omega_1 g/\omega_{2,\infty}^2 + h}{1 - \omega_{1,\infty}^2/\omega_{2,\infty}^2}$$

$$m_2 = \frac{g + \omega_{1,\infty}^2 h/\omega_{-1}\omega_1}{1 - \omega_{1,\infty}^2/\omega_{2,\infty}^2},$$

where

$$\omega_{1,\infty} = 2\pi \cdot 3300;\ \omega_{2,\infty} = 2\pi \cdot 8300$$

$$g = \left[\left(1 - \frac{\omega_{1,\infty}^2}{\omega_1^2}\right)\left(1 - \frac{\omega_{1,\infty}^2}{\omega_2^2}\right)\right]^{1/2}$$

$$h = \left[\left(1 - \frac{\omega_1^2}{\omega_{2,\infty}^2}\right)\left(1 - \frac{\omega_2^2}{\omega_{2,\infty}^2}\right)\right]^{1/2}.$$

The parameter for the M-type section IV_1, M is

$$M = O \cdot 6^*.$$

Instead of making the further calculations with the other formulas given on pp. 41 and 42 (l.c.) we prefer to go back to the formulas (15), (19), (24) and (25) (l.c.).

1. For the constant k-type VII with $z_{a1}/2 = R\lambda/(\omega_1 - \omega_{-1})$, $z_{a2}/2 = \omega_{-1}\omega_1 R/(\omega_1 - \omega_{-1})\lambda$ (p. 42, l.c.) we get

$$Z_{1,\mathrm{VII}} = \tfrac{1}{2}z_{1k} = \tfrac{1}{2}z_{a1} + \tfrac{1}{2}z_{a2} = \frac{R(\lambda^2 + \omega_{-1}\omega_1)}{(\omega_1 - \omega_{-1})\lambda}$$

$$Z_{2,\mathrm{VII}} = \tfrac{1}{2}z_{1k} + 2z_{2k} = \tfrac{1}{2}z_{1k} + \frac{2R^2}{z_{1k}} = \frac{R(\lambda^2 + \omega_{-1}{}^2)(\lambda^2 + \omega_1{}^2)}{(\omega_1 - \omega_{-1})\lambda(\lambda^2 + \omega_{-1}\omega_1)}$$

and

$$\left(\frac{Z_2}{Z_{1,\mathrm{VII}}}\right)^{1/2} = \frac{[(\lambda^2 + \omega_{-1}{}^2)(\lambda^2 + \omega_1{}^2)]^{1/2}}{\lambda^2 + \omega_{-1}\omega_1} \text{ (attenuation class 2).}$$

2. For IV_1 we have

$$Z_{1,\mathrm{IV1}} = \tfrac{1}{2}z_{11} = \tfrac{1}{2}m_1 z_{a1} + \tfrac{1}{2}m_2 z_{a2} = \frac{Rm_1(\lambda^2 + m_2\omega_{-1}\omega_1/m_1)}{(\omega_1 - \omega_{-1})\lambda}$$

and according to (15), l.c.

$$z_{11}z_{21} = R^2 + \left[\frac{R(\lambda^2 + \omega_{-1}\omega_1}{(\omega_1 - \omega_{-1})\lambda}\right]^2 - \left[\frac{R(m_1\lambda^2 + m_2\omega_{-1}\omega_1)}{(\omega_1 - \omega_{-1})\lambda}\right]^2$$

$$Z_{2,\mathrm{IV}_1} = \tfrac{1}{2}z_{11} + 2z_{21} = \frac{R(\lambda^2 + \omega_{-1}{}^2)(\lambda^2 + \omega_1{}^2)}{(\omega_1 - \omega_{-1})m_1\lambda(\lambda^2 + m_2\omega_{-1}\omega_1/m_1)}$$

$$\left(\frac{Z_2}{Z_{1,\mathrm{IV}_1}}\right)^{1/2} = \frac{[(\lambda^2 + \omega_{-1}{}^2)(\lambda^2 + \omega_1{}^2)]^{1/2}}{m_1(\lambda^2 + m_2\omega_{-1}\omega_1/m_1)} \text{ (attenuation class 2).}$$

3. For $IV_{1,M}$ there follows according to (19) and (25), l.c.

$$Z_{2,\mathrm{IV}_{1,M}} = 2z_{21} = \frac{1 - M^2}{2M}z_{1k} + \frac{2R^2}{Mz_{1k}}$$

$$= \frac{R[(1 - M^2)\lambda^4 + (\omega_{-1}{}^2 - 2M^2\omega_{-1}\omega_1 + \omega_1{}^2)\lambda^2 + (1 - M^2)\omega_{-1}{}^2\omega_1{}^2]}{(\omega_1 - \omega_{-1})M\lambda(\lambda^2 + \omega_{-1}\omega_1)}$$

* We write M for Zobel's m in order to avoid a confusion with our m in Table I.

$$z_{11} = Mz_{1k}$$

$$Z_{1,\mathrm{IV}_{1,M}} = \frac{2Mz_{21}z_{1k}}{2Z_{2,\mathrm{IV}_{1,M}} + Mz_{1k}} = \frac{2Mz_{21}z_{1k}}{4z_{21} + Mz_{1k}}$$

$$= \frac{MR(\lambda^2+\omega_{-1}\omega_1)\left\{(1-M^2)\lambda^4+(\omega_{-1}^2-2M^2\omega_{-1}\omega_1+\omega_1^2)\lambda^2+(1-M^2)\omega_{-1}^2\omega_1^2\right\}}{(\omega_1-\omega_{-1})\lambda(\lambda^2+\omega_{-1}^2)(\lambda^2+\omega_1^2)}$$

$$\left(\frac{Z_2}{Z_{1,\mathrm{IV}_{1,M}}}\right)^{1/2} = \frac{[(\lambda^2+\omega_{-1}^2)(\lambda^2+\omega_1^2)]^{1/2}}{M(\lambda^2+\omega_{-1}\omega_1)} \text{ (attenuation class 2)}$$

$$(Z_1Z_2)^{1/2} = (Z_1Z_2)^{1/2}{}_{\mathrm{IV}_{1,M}}$$

$$= \frac{R[(1-M^2)\lambda^4+(\omega_{-1}^2-2M^2\omega_{-1}\omega_1+\omega_1^2)\lambda^2+(1-M^2)\omega_{-1}^2\omega_1^2]}{(\omega_1-\omega_{-1})\lambda[(\lambda^2+\omega_{-1}^2)(\lambda^2+\omega_1^2)]^{1/2}}.$$

This is the image impedance of the whole chain and belongs to class *d*. It follows from the additive rule for the attenuation class numbers that the attenuation class of the whole chain is 6, or the class of the band-pass filter of Fig. 13 is found to be 6*d*.

Rule (15) gives for the whole chain of Fig. 13,

$$\left(\frac{Z_1}{Z_2}\right)^{1/2} = \frac{\lambda^6(1+m_1)+A\lambda^4+B\lambda^2+C}{(\lambda^2+\omega_{-1}^2)^{1/2}(\lambda^2+\omega_1^2)^{1/2}[\lambda^4(1+m_1)+D\lambda^2+E]}$$

where

$$A = \omega_{-1}^2+\omega_{-1}\omega_1+\omega_1^2+m_2\omega_{-1}\omega_1+\frac{m_1}{1+M}(\omega_{-1}^2+2M\omega_{-1}\omega_1+\omega_1^2)$$

$$B = \omega_{-1}\omega_1(\omega_{-1}^2+\omega_{-1}\omega_1+\omega_1^2)+m_1\omega_{-1}^2\omega_1^2+\frac{m_2}{1+M}\omega_{-1}\omega_1(\omega_{-1}^2+2M\omega_{-1}\omega_1+\omega_1^2)$$

$$C = (1+m_2)\omega_{-1}^3\omega_1^3$$

$$D = (m_1+m_2)\omega_{-1}\omega_1+\frac{1}{1+M}(\omega_{-1}^2+2M\omega_{-1}\omega_1+\omega_1^2)$$

$$E = (1+m_2)\omega_{-1}^2\omega_1^2.$$

MULTIPLE PEAKS IN SUPERSONIC INTERFEROMETRY*

By W. D. Hershberger
Signal Corps Laboratories, Fort Monmouth, New Jersey

(Received February 15, 1932)

Abstract

An improved circuit for use in supersonic interferometry in gases is described in which regeneration is under control and by the use of which the amplitude of the exciting e.m.f. may be varied between wide limits. Multiple peaks in the plate current maxima are observed, and several causes for such peaks, particularly the use of crystals having "parasitic" frequencies, are pointed out.

Introduction

THE present experimental work follows lines similar to those of previously published work in supersonic interferometry.[1] However, some observations on the sound radiated by crystals having "parasitic" frequencies are reported and a more complete analysis of the basis of interferometer action is given to account for an observed complexity in the maxima in the plate current curves.

Apparatus

An improved circuit was employed which possesses the advantage that the electrical regeneration is under control. The circuit diagram appears in Fig. 1. Control of regeneration is effected by means of the potentiometer marked P instead of permitting regeneration to depend upon such factors as the length and disposition of the leads to the crystal, the capacity between the electrodes in the tube, and other usual sources of stray coupling. The circuit bears some resemblance to the Hartley type but is adapted to crystal control. When the potentiometer contact is at the point of highest a.c. potential with respect to the grounded power supply, the amplitude of the exciting e.m.f. attains a maximum value and the crystal is least effective in stabilizing frequency, but as the slider is moved in the direction of lower a.c. potential with respect to ground, the amplitude of the e.m.f. across the crystal decreases, frequency stability increases, and the reaction of changes in the length of the gas column on the plate current becomes more and more violent. Thus the amplitude of the 100th peak in plate current using a 497.5 kc, crystal and a 112-A tube was 8 milliamperes. When regeneration is reduced to a bare

* Published with permission of the War Department.

[1] G. W. Pierce, Am. Acad. Proc. **60**, 271 (1925); W. H. Pielemeier, Phys. Rev. **34**, 1184 (1929); **36**, 1005 (1930); G. E. Thompson, Phys. Rev. **36**, 77 (1930); C. D. Reid, Phys. Rev. **35**, 814 (1930); **37**, 1147 (1931); Hubbard and Loomis, Phil. Mag. **5**, 1177 (1928); Loomis and Hubbard, J.O.S.A. **17**, 295 (1928); J. C. Hubbard, Phys. Rev. **35**, 1442 (1930); **36**, 1668 (1930); Klein and Hershberger, Phys. Rev. **37**, 760 (1931).

minimum, the crystal is sufficiently loaded by the gas column that it refuses to oscillate for any reflector positions excepting certain critical ones with a spacing of a half wave-length between neighboring positions. This results in a flat-topped plate current curve with sharp depressions.

The shifts in frequency incident to changes in the length of the gas column were measured by observations on the beat note between the driving circuit frequency and that of an independent monitoring circuit. The calibration of this latter circuit was effected by the use of auxiliary quartz crystals.

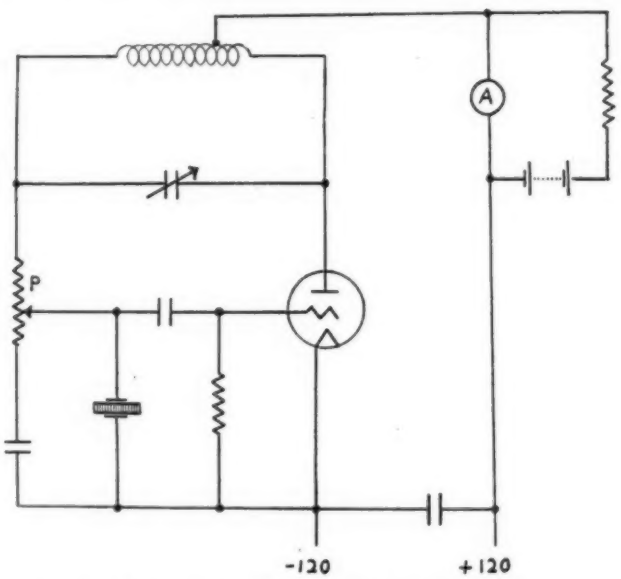

Fig. 1. Circuit in which regeneration is under control.

An interferometer of simple and inexpensive design was used in some of this work. A Starrett $2M$ micrometer was modified by replacing the anvil by a stem, integral with a shallow cup which contains the crystal. A brass reflecting disk was machined to permit a friction fit on the spindle of the micrometer. The face of the reflector and the bottom of the crystal cup were kept accurately parallel. The distance of the reflector from the radiating face of the crystal could be varied from 1.000 cm to 3.200 cm. A diagram of this instrument is shown in Fig. 2.

The crystals used in the study were of various frequencies ranging from 112 kc to 575 kc. Several crystals were kindly loaned by Professor George Winchester of Rutgers University while a number of others belonged to these Laboratories and to the writer.

Results

Multiple peaks of two distinct types were observed. One type was associated with a relatively large frequency shift as the length of the gas column was varied, while in the second type the frequency shift is extremely small. That is, it is of the same order of magnitude as that first reported by Pierce.

One X-cut crystal displayed two resonance frequencies of 571.1 kc and 572.3 kc. By suitable adjustment of the tuning condenser and the regeneration it was possible to shift from one frequency to the other—a change of 1.2 kc—merely by varying the length of the air column. Very pronounced double peaks resulted. The circuit adjustments when using this crystal were the controlling factor in the production of peaks of the first type. Crystals having "parasitic" frequencies relatively close together may be made to display this effect.

The second type of multiple peak was observed using the previously mentioned X-cut crystal for circuit adjustments which did not permit an appreciable frequency shift. The same type of peak was observed when using

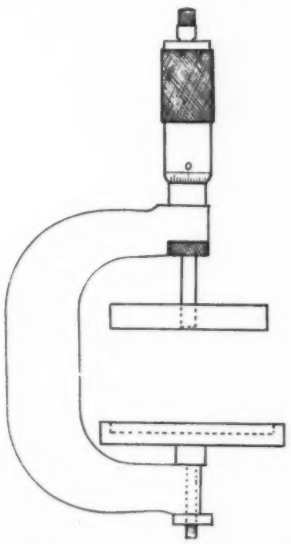

Fig. 2. Diagram of interferometer.

a Y-cut crystal having resonance frequencies at 370.3 kc, 404.8 kc, 412.3 kc, and 430.0 kc. The maxima in the plate current curves taken at 370.3 kc—at which frequency the crystal was a powerful oscillator—were singly-peaked but at the other frequencies multiply-peaked. Oscillations at these latter frequencies were relatively feeble. All of these observations employed sound radiation from the same face of the crystal. The frequency shifts accompanying changes in the reflector position were not large enough to account for the multiple peaks. When employing another X-cut crystal with resonance frequencies of 497.5 kc and 510.0 kc, the maxima showed no evidence of multiple peaks. From the present observations no general rules governing the appearance of multiply-peaked maxima can be laid down which may be expected to hold in every case. However, in general, X-cut crystals with resonance frequencies relatively far apart and at the same time powerful oscillators gave singly-peaked maxima and appeared to be especially suitable for interferometric work. On the other hand, some Y-cut crystals and all crystals which were feeble oscillators were more erratic in their behavior and

were prolific in the production of multiply-peaked maxima. Typical curves are shown in Fig. 3. Curve (a) was taken employing circuit adjustments which did not permit an appreciable frequency shift; for slightly different circuit adjustments the frequency shifts from 571.1 kc and 572.3 kc as the column length is varied. Curve (b) was taken when employing a 497.5 kc crystal. The change in plate current here was 10 m.a.

Multiple peaks resembling those here reported have earlier been described by Pielemeier.[2] He assumed two velocities of sound to explain his observations:—a low velocity for low intensity sound to account for minor peaks and a higher velocity for high intensity sound to account for major peaks. This hypothesis was found to be inadequate to explain multiple peaks of both types reported here. However, there may be still other types. No shift in the position of the major peaks was detected as the amplitude of the exciting

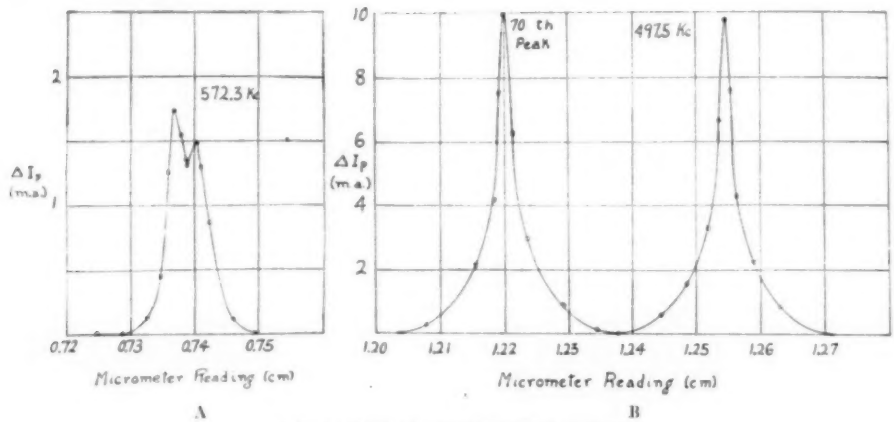

Fig. 3. Typical experimental curves.

e.m.f. was varied over a wide range while holding frequency constant. This was effected by introducing wide variations in the plate voltages employed as well as by changes in the regeneration and using a tube of the 210 type. A shift in the position of the major peaks calculated on the basis of the velocity dependence on intensity which Pielemeier assumes would have been detected very readily. Also the multiple peaks persisted at both low and high amplitude of the exciting e.m.f. In some of the experiments at very low amplitudes, multiple peaks were especially pronounced. None of the observations led to abnormal velocities of sound when corresponding points on successive maxima were used in calculating wave-lengths. The monitoring circuit proved imdispensable especially when working with crystals with a number of resonance frequencies.

Theory

A closer examination of the basis of interferometer action reveals a reasonable explanation for the appearance of multiply-peaked plate current maxima. The first type is unmistakably due to a shift in frequency as the

[2] W. H. Pielemeier, Phys. Rev. **38,** 1236 (1931).

length of the air column is varied. However, the second type can not be explained on this basis. Rather these are to be viewed as arising from the fact that the crystal is capable of oscillating at any one of a number of frequencies with the phase relations at the radiating surface depending on the particular mode of vibration. Previous theoretical work[3] has not predicted the existence of multiple peaks. The writer has shown, assuming a single frequency for the crystal, that the mechanical impedance of the equivalent piston of area A coupled to a fluid column of variable length 1, is:

$$z = z_0 + z(l) \tag{1}$$

where

$$z_0 = j(mw - s/w) + r \tag{2}$$

$$z(l) = Ac\rho\left[\frac{e^{4\alpha l} - 1}{e^{4\alpha l} + 2e^{2\alpha l}\cos\phi + 1} + j\frac{2e^{2\alpha l}\sin\phi}{e^{4\alpha l} + 2e^{2\alpha l}\cos\phi + 1}\right] \tag{3}$$

and

$$\phi = \frac{2\omega l}{c} - \pi. \tag{4}$$

Z_0 is the mechanical impedance of the piston itself considered apart from sound radiation, while $Z(l)$[4] is the impedance due to the reaction of the fluid column on the piston. c is the velocity of sound in the fluid, ρ its density, and α the amplitude damping constant per unit path length. If the absolute value of $Z(l)$ is plotted against l, we obtain singly-peaked maxima. But it is particularly to be noted that we have assumed that all of a surface of area A is radiating sound in phase. This assumption is not justified for crystals capable of oscillating at any one of a number of frequencies. Then the mechanical impedance at the driving point is a more complicated function of frequency, the crystal constants, and $Z(l)$. This function must remain unknown in the absence of specific knowledge concerning the modes of vibration of the quartz plate and the coupling between modes. The electrical circuit elements[5] used as analogies to account for some aspects of the behavior of crystals with many degrees of freedom if combined with impedances of the type $Z(l)$ permit an explanation for multiply-peaked maxima. Impedances of this latter type must enter the simultaneous force equations used in discussing vibration problems involving several degrees of freedom as many times as the number of modes for which we have appropriate gas column reactions. In short, we look for an explanation of the observed complexity in the plate current curves to the known complexity in the motion of the crystal itself. Methods based on supersonic interferometry may prove useful in studying the modes of vibration of such crystals.

[3] W. D. Hershberger, Jour. Acoustical Soc. of Amer. **3,** 263 (1931); J. C. Hubbard, Phys. Rev. **38,** 1011 (1931).

[4] It may be pointed out that $R(1)/Ac\rho$ of equation (3) above is identical with Professor Hubbard's P of his equation (14) and $X(1)/Ac\rho$ with his Q of equation (15).

[5] F. R. Lack, Bell System Tech. Jour. **8,** 515 (1929).

INFLUENCE OF HYDROGEN ION CONCENTRATION ON THE CRYSTAL STRUCTURE OF ELECTRODEPOSITED COBALT

By H. Kersten
Department of Physics, University of Cincinnati

(Received January 14, 1932)

Abstract

Cobalt crystallizes in the cubic and hexagonal systems. This paper shows by means of x-ray diffraction analysis, that cobalt electrodeposited from high *pH* (alkaline) sulphate solutions has a hexagonal structure which changes into a mixture of hexagonal and cubic, as the *pH* is decreased.

HULL[1] has shown that cobalt has hexagonal close-packed as well as face centered cubic lattices, and that both types of lattice are obtained by rapid electrolysis of the sulphate. The purpose of this investigation was to determine whether the hydrogen ion concentration of the electrolyte influences the structure of the electrodeposited cobalt.

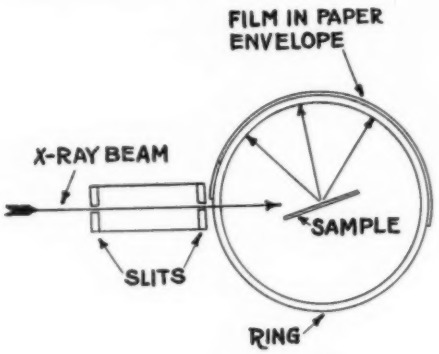

Fig. 1. Schematic diagram of the camera.

Experimental

Cobalt was deposited from an electrolyte having the following composition:[2] Cobalt sulphate ($CoSO_4:7H_2O$), 504 gm/liter; sodium chloride (NaCl), 17 gm/liter; boric acid (H_3BO_3), 45 gm/liter, at 30°C and 1 amp/dm² with cast cobalt anodes (98–99 percent pure) and stainless steel cathodes. The hydrogen ion concentration was adjusted with H_2SO_4 and NH_4OH and measured with a quinhydrone electrode.

The x-rays were supplied by a chromium target gas-type tube[3] operated

[1] Hull, Phys. Rev. **17**, 571 (1921).

[2] Blum and Hogaboom, Principles of Electroplating and Electroforming, 2nd ed., p. 280, 1930.

[3] Kersten, Rev. Sci. Inst. **3**, 145 (1932).

at 20 kv and 40 m. a. The camera is illustrated schematically in Fig. 1. The diameter of the ring around which the film was wrapped was such that 0.025 inch on the film corresponded to 1°. The black line at the right end of each photograph in Fig. 2 was made by the main beam before the sample was put in place.

Theoretical

The positions of the diffracted images of the main x-ray beam, plotted from the data given by Hull,[1] are shown for both lattice types in the lower part of Fig. 2. All the theoretically possible lines do not appear in the photographs but it is evident that the 200 $K\alpha$ line, indicated by an arrow, is the best

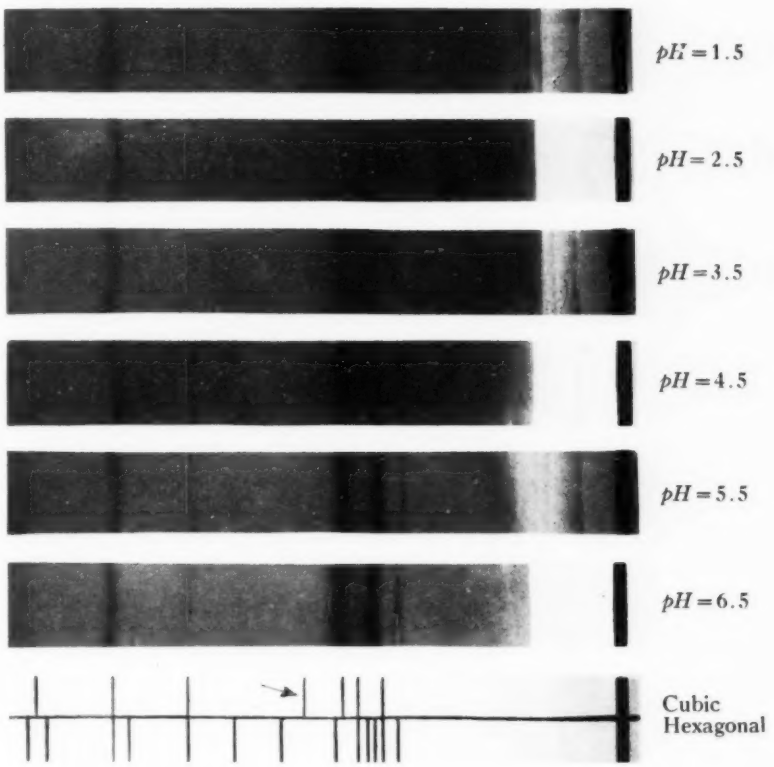

Fig. 2. Influence of pH on the crystal structure of electrodeposited cobalt.

one to use as a test for the presence of cubic cobalt since it is one of the strongest lines appearing in the photographs when a large percentage of cubic crystals are present in the cobalt.

Results

Fig. 2 indicates that hexagonal cobalt is deposited at high values of pH from the electrolyte used and that the structure gradually changes into a mixture of hexagonal and cubic as the pH is decreased.

INFLUENCE OF TEMPERATURE ON THE CRYSTAL STRUCTURE OF ELECTRODEPOSITED ANTIMONY

By H. Kersten
Department of Physics, University of Cincinnati

(Received February 25, 1932)

Abstract

It is shown by x-ray diffraction methods that "explosive" antimony is amorphous before exploding and crystalline afterward. It is also shown that the crystal structure of electrodeposited antimony is influenced by the temperature of the plating bath.

Introduction

Gore[1] found that antimony, electrodeposited from a cold solution of antimony trichloride in dilute hydrochloric acid, has the property of "exploding" when scratched and that deposits from a hot bath do not have this

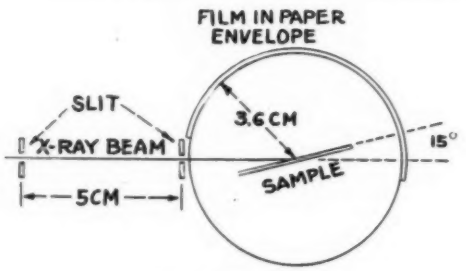

Fig. 1. Camera.

property. The investigation described below was intended to determine whether the explosive property is associated with a particular crystal structure and to find the influence of bath temperature in general, on the structure of electrodeposited antimony.

Experimental

The plating bath used had the following composition: Antimony trichloride ($SbCl_3$), 200 gm; hydrochloric acid (HCl), 300 cc; water, to make, 1 liter. Cast antimony anodes and nickel-plated brass cathodes were used. The cathodes were nickel-plated to prevent the antimony from depositing by "immersion" on the brass and possibly alloying with it. The current density in all the experiments was 3 amp. cm^2.

The x-rays were supplied by a special gas tube[2] having a copper target and operated at 40 m. a. and 20 kv. The camera for holding the film and sample is

[1] Gore, Phil Mag. (4) **9**, 73 (1855).
[2] Kersten, Rev. Sci. Inst. **3**, 145 (1932).

illustrated in Fig. 1. The antimony was plated thick enough to prevent the diffraction lines from the nickel and brass from appearing on the film. This made it unnecessary to strip the antimony from the cathode when taking a picture.

Results

Samples plated at 5° intervals from 0°C to 100°C are shown in Fig. 2, 3, and 4. It is evident that the electrodeposited antimony has an amorphous

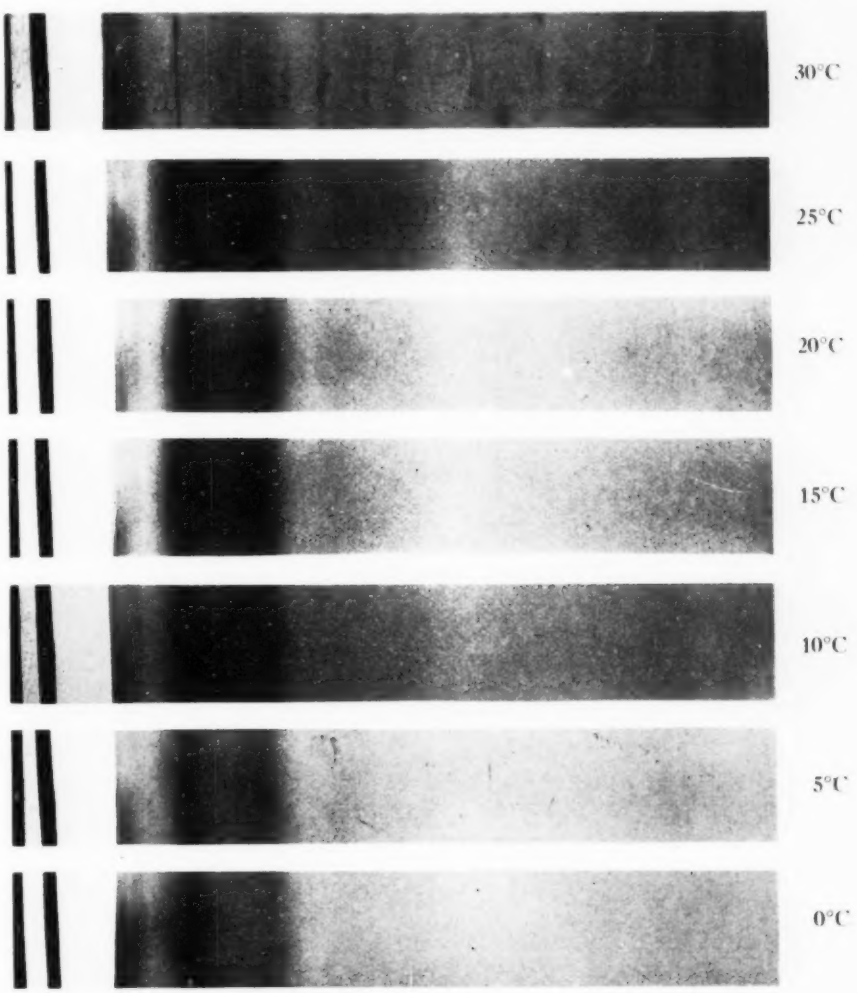

Fig. 2. Diffraction photographs of antimony.

structure when electroplated at temperatures below 25°C and that it changes from amorphous to crystalline between 25° and 30°. Photographs taken for samples plated in this range, but not shown in the figures, showed that the change was a gradual one. The structure from 30° to 70° is nearly constant. Above 70° there are changes in the position and intensity of lines which indicate changes in structure.

When a cathode plated with a thick layer of the amorphous modification is scratched with the point of a knife it gives off a white vapor accompanied by a hissing sound and a rise in temperature. In order to determine whether this change is accompanied by a change in structure, a sample of the amorphous variety was clamped in the camera and the photograph shown in the middle part of Fig. 5 taken. Without removing the sample from the camera,

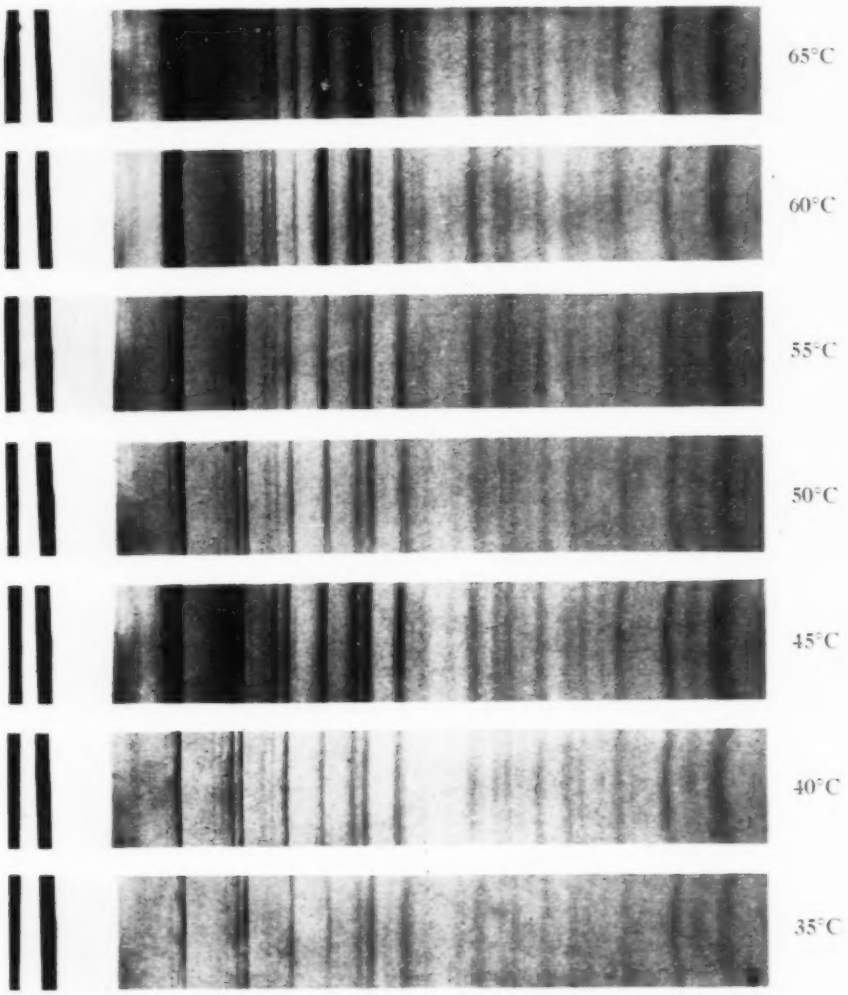

Fig. 3. Diffraction photographs of antimony.

its surface was scratched to cause the explosion and another photograph, shown in the upper part of Fig. 5, was taken. These show that the antimony had changed from an amorphous to a crystalline form during the explosion. The lower part of Fig. 5 shows a photograph of chemically pure antimony. A comparison of the upper and lower photographs shows that the structure of the electrolytic amorphous antimony after being exploded, is the same as that of pure antimony.

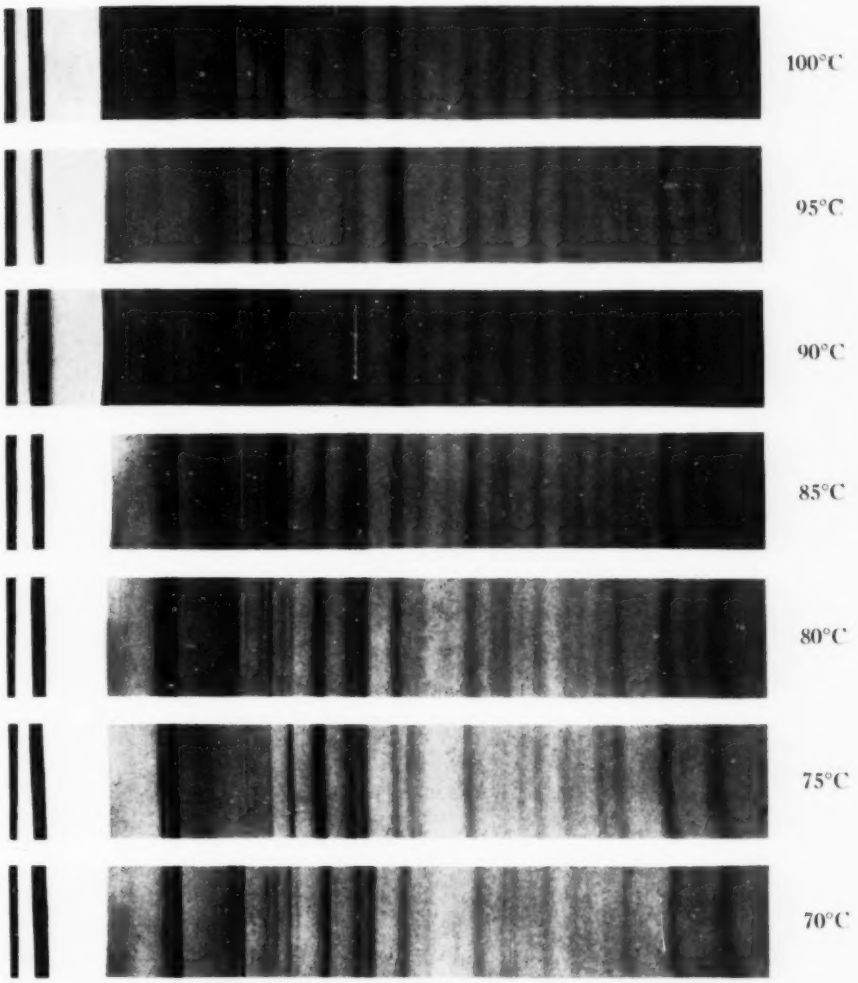

Fig. 4. Diffraction photographs of antimony.

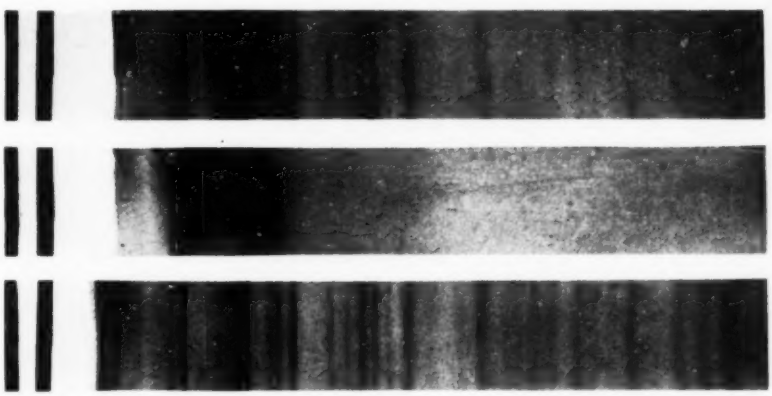

Fig. 5. Upper: Antimony after exploding; middle: Antimony before exploding; lower: Pure antimony.

CATHODE SPUTTERING, A COMMERCIAL APPLICATION

By H. F. Fruth
Western Electric, Chicago

(Received February 15, 1932)

Abstract

The theory of cathode sputtering with the advantages and limitations in the application of this process is discussed, followed by a description of the commercial equipment and methods for applying gold electrode surfaces to diaphragms of certain types of microphones. By proper design of the vacuum chamber and the inside parts, a fairly uniform discharge current density and a uniform deposit is obtained. A constant sputtering rate is produced by the use of a bleeder valve which maintains a proper residual pressure. Adherence and continuity are obtained by the use of a special cleaning process. An extensive bibliography on cathode sputtering is included.

Introduction

ALTHOUGH the process of electrostatic deposition of metals by high voltages in a partial vacuum, commonly known as cathode sputtering, has been known for more than a half century, it has heretofore found but little commercial application. Rather extensive use of it has, however, been made in physics research laboratories for such purposes as the production of highly reflecting surfaces on mirrors and prisms, for spectrometers, interferometers, etc., and the making of extremely thin metal films for fundamental studies in atomic structure and electron theory. Sputtering has also been used in the manufacture of very fine conducting quartz fibres for suspensions in sensitive instruments such as quadrant and string electrometers, galvanometers, and electrocardiographs, and, to some extent, for etching certain metals.

In the following paragraphs, it is intended to give a brief explanation of the process and a description of a commercial application in the production of diaphragms for certain microphones.

Theory

Some fifty years ago, various investigators working on high voltage discharges in vacuo discovered that disintegration of the cathode occurs for nearly all metals and that the removed metal is deposited in a very fine state of subdivision on nearby objects. Various theories have been formulated as to the mechanism of this phenomenon. Some investigators have attempted to explain it by stating that it is a type of electrical evaporation where the electrical potential is analogous to the temperature potential in ordinary thermal evaporation. More plausible theories, recently advanced, are that the metal atoms or particles leave the cathode as a result of positive ion bombardment, either as a result of single impacts or by cumulative action. This does not seem to hold for all metals and residual gases, and it has been suggested that

in some cases sputtering is due to the absorption of radiation produced when the gas ions are stopped at the cathode.

There has also been a great deal of controversy as to the structure of the sputtered film. Most of the more recent evidence indicates that the films are of a crystalline nature.

Advantages of Sputtering over Other Methods of Metallizing

In general, cathode sputtering cannot compete with the more common processes of metallic deposition, but in many special cases it can be employed where other methods are inapplicable. This is especially true (a) when metals are to be deposited upon non-conductors, (b) when the surfaces to be metallized would be injured by contact with chemical solutions or high temperature, (c) when either a very thin continuous metal coat or a very smooth, highly reflecting coat is desired, (d) when metals are to be deposited that are very difficult to deposit in any other manner, such as, silicon, tellurium, or

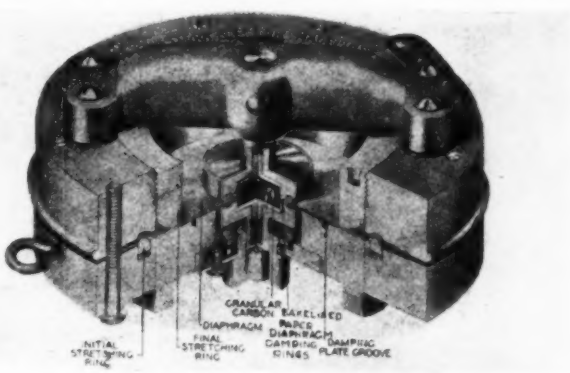

Fig. 1. Carbon broadcasting transmitter.

selenium, (e) when a metal is to be deposited upon another metal far removed from it in the electrochemical series, such as, gold or platinum upon aluminum or magnesium. In such cases, electrolytic corrosion is very apt to occur in the presence of moisture or traces of electrolyte.

Commercial Application

Cathode sputtering has been used commercially by the Western Electric Company with excellent results in the manufacture of diaphragms for carbon broadcasting transmitters, illustrated in Fig. 1. This type of microphone consists essentially of a tightly stretched duralumin diaphragm located between two chambers containing granular carbon. The manner in which the double button type of microphone is connected to the amplifiers of the broadcasting, public address, or sound picture system is schematically shown in Fig. 2. Those parts of the diaphragm which come in contact with the carbon are covered with gold, a metal which has been found to be ideally suited for microphonic purposes. It is exceedingly important that the gold coating be

continuous and remain adherent, for if the carbon should make contact in a number of places with the duralumin, it is quite likely that the microphone would be noisy and therefore unfit for use.

When the gold spots were formed by the plating process previously used, considerable trouble was experienced because of such defects as blisters, peeling, and pin holes probably due largely to electrolytic corrosion. The substitution of the sputtering process has practically eliminated such defects. Some additional advantage over electrolytic plating is obtained in that a thinner

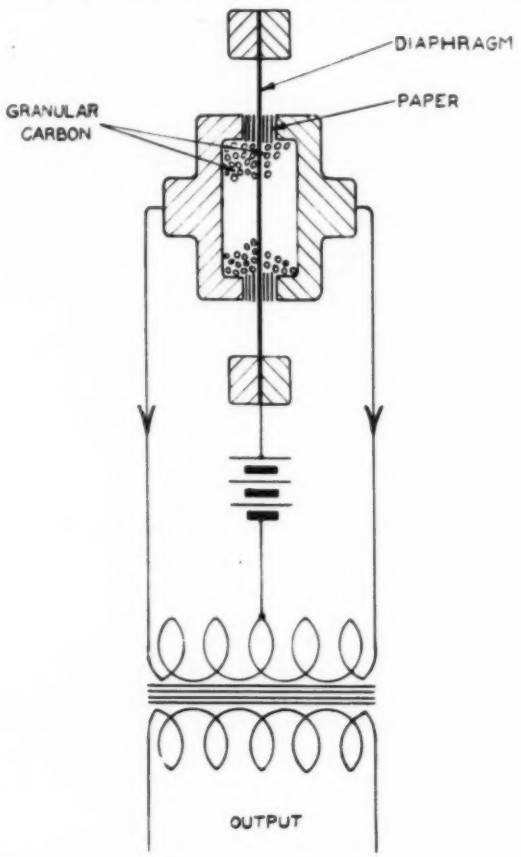

Fig. 2. Diagram of connections between double button type of microphone and amplifier.

continuous film can be deposited, making it possible to stretch the diaphragm to a higher natural frequency with less tension. This decrease in tension has resulted in a smaller frequency loss due to fatigue. The introduction of the sputtering process has, therefore, not only improved the quality of the microphone, but lengthened its life.

Equipment and Method

In order to apply a satisfactory gold electrode surface to the duralumin diaphragm by cathode sputtering commercially, it was necessary to develop

a special cleaning process for the diaphragms and suitable equipment with multiple sputtering electrodes.

To insure proper continuity and adherence of the gold to the duralumin, the glossy roll finish on the duralumin is broken up by means of a brass wire scratch brush to give a matte surface. The diaphragms are then scrubbed in acetone, rinsed in ether, and rubbed dry with filter paper to remove all traces of oil or grease. Great care is then taken to keep the spot to be coated free from all contaminations before sputtering.

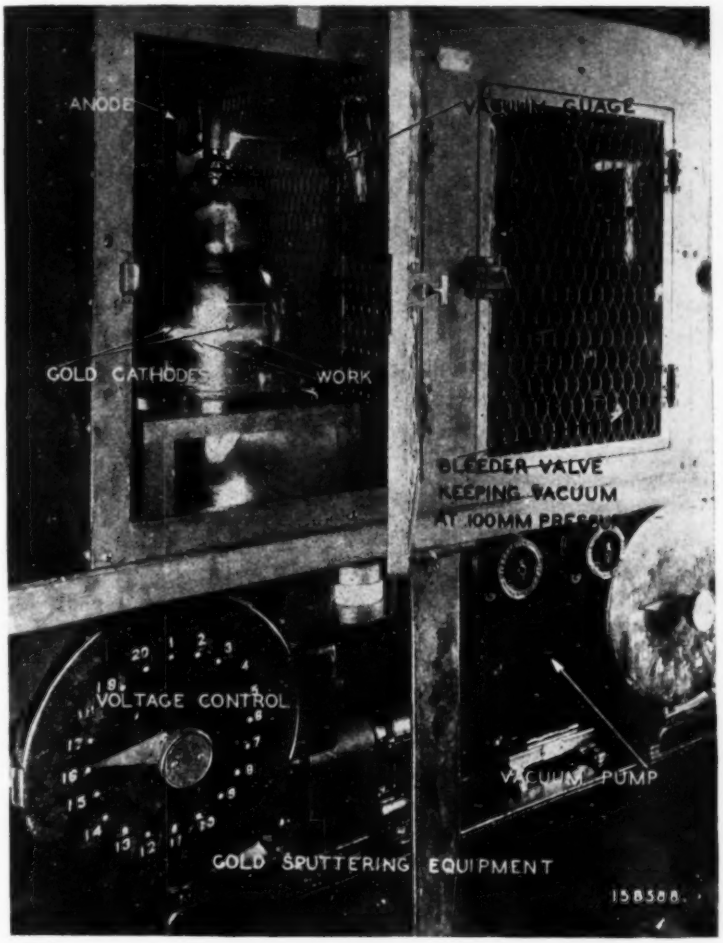

Fig. 3. Commercial sputtering unit.

The unit developed for commercial sputtering is shown in operation in Fig. 3. A two-gallon bell jar fitted upon a heavy ground Pyrex glass plate contains six glass-covered aluminum cathodes, the exposed parts of which are fitted with renewable gold disks (Fig. 4). An adjustable aluminum stand to hold the diaphragm fixture and shield is placed upon the plate directly below

the cathodes. The shield covers all of each diaphragm except a central circular spot about 3/4″ in diameter. A vacuum is produced by means of a Megavac pump connected to the bell jar by 5/8″ Pyrex glass tubing, which is made as short as possible to facilitate rapid pumping. In the vacuum system are a small rotating McLeod gauge, a trap, a bleeder valve, and an anode. The anode consists of a 20-turn, 1/2″ helix of No. 14 aluminum wire sealed into a one-inch tube at some distance from the bell jar. The end of the helix is

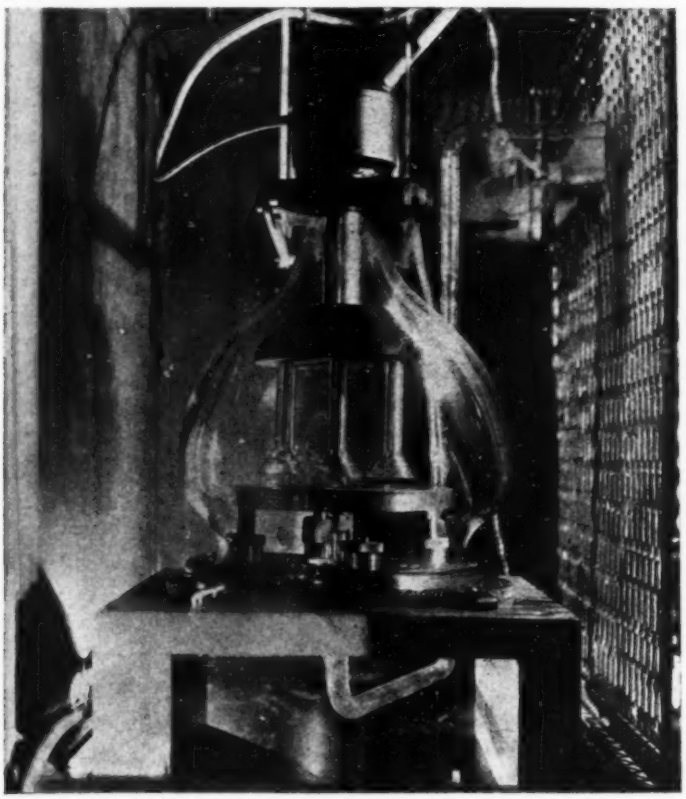

Fig. 4. Cathodic sputtering unit for microphone diaphragms.

shielded against high current density by a flanged piece of glass tubing. This type of anode has a life of at least two thousand hours.

In order to maintain a constant residual gas pressure, the pump is operated continuously and air is allowed to leak in slowly through the bleeder valve which is located near the pump. This practice was found necessary in order to overcome variations in pressure due to the early evolution of gases and the later cleanup usually accompanying electrical discharges in vacuo. The valve is of rugged constructuon as shown in Fig. 5 and consists of a standard No. 0 taper pin about 2 1/2 inches long, very closely lapped into a bronze bushing. A pressure of 0.100 ± 0.005 mm is readily maintained by this method. After a new charge has been placed in the bell jar, the bleeder valve

is temporarily cut off by closing a stop cock so that the required vacuum can be more quickly obtained. By this means, sputtering can be started in about four minutes after the bell jar has been placed in position.

The discharge is produced by means of a 1/2 k-v-a transformer which steps up the voltage from 110 to 10,000 volts and is regulated by means of a rheostat in the primary. Safety for the operator is insured by having all high tension leads and terminals enclosed in an expanded metal cage. The door operates a switch in the primary circuit, so that as the door is opened, the circuit is broken.

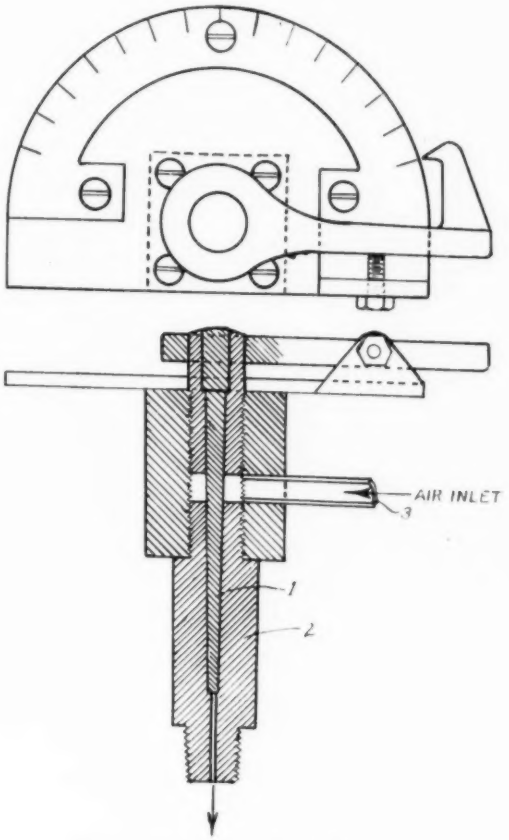

Fig. 5. Bleeder valve.

In order that equal amounts of metal might be evenly deposited on each of the six diaphragms, it was necessary to determine experimentally the relative spacing of the electrodes, and the shape and the size of the shields, stands, and bell jar which would give a uniform current distribution. The best results were obtained by placing the diaphragms 5/8″ from the cathodes, which are arranged symmetrically, by maintaining the pressure at 0.10 ± 0.05 mm, by having a 1/2″ hole in the center of the diaphragm holder and shield, and a one inch clearance between their outer rims and the bell jar.

In commercial practice a single operator runs three units simultaneously and produces about 90 broadcasting microphone diaphragms a day. These diaphragms require a gold spot on both sides which necessitates breaking the vacuum and turning the parts over. The time not occupied in loading the fixtures and operating the equipment is taken up in preparing diaphragms for sputtering, marking, and packing the finished product. It requires about forty minutes to deposit from three to five mg of gold per spot. Although this deposit is not more than 0.001 mm thick, it is very continuous, adherent, and altogether suitable as a microphonic electrode surface. (Fig. 6.)

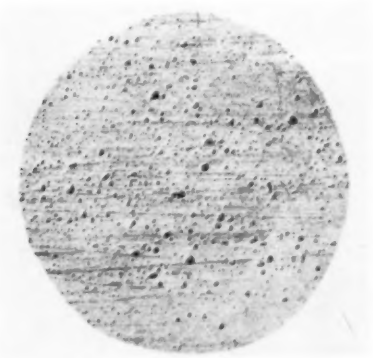

Fig. 6a. Representative gold plated diaphragm surface after 3 months service showing blisters.

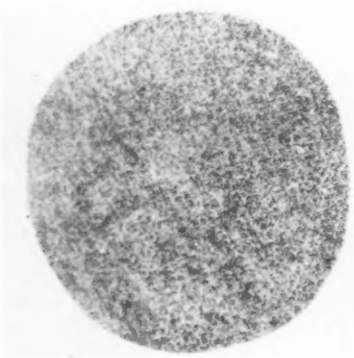

Fig. 6b. Representative gold sputtered diaphragm surface after 3 months service showing no blisters.

This equipment has been in successful operation for a number of years in the production of microphone diaphragms.

Bibliography

W. R. Grove, Phil. Trans. **1,** (1852).

J. P. Gassiot, Proc. Roy. Soc. **9,** 146 (1858).

Plucker, Pogg. Ann. **103,** 90; **104,** 116; **105,** 70 (1858).

Gassiot, Rep. of Brit. Assoc. **2,** 38 (1861); Fortschr. der Phys. **II,** 505 (1862).

Wright, Amer. Journ. of Sc. (Silliman) **13,** 49; **14,** 169 (1877).

Puluj, Wien. Ber. **81,** II, 864 (1880).

Fr. Wachter, Wied. Ann. **17,** 903 (1882).

E. Wiedemann, Wied. Ann. **20,** 756 (1883).

Hittorf, Wied. Ann. **21,** 126 (1884).

A. Kundt, Wied. Ann. **27,** 59 (1886) and B. Dessau, Wied. Ann. **29,** 353 (1886).

W. Crookes, Proc. Roy. Soc. **50,** 88 (1891).

Moser, Wied. Ann. **42,** 639 (1891).

G. Granquist, Ofvers. Svensk. Vetensk. Ak. Forh. **54,** 575 (1897).

G. Granquist, Ofvers. Svensk, Vetensk. Ak. Forh. **55,** 709 (1898).

Wehnelt. Wied. Ann. **67,** 425 (1899).

A. C. Longdon, Phys. Rev. I, **5,** 11, 40, and 84 (1900).

J. Patterson, Cam. Phil. Soc. Proc. **V 11,** 118 (1901).

J. J. Thompson, Cam. Phil. Soc. Proc. **V 11,** 120 (1901).

Goldstein, Verh. d. D. Phys. Ges. 237 (1902).

W. Wien, Phys. Zeits. **3,** 440 (1902).

G. C. Schmidt, Ann. d. Physik **9,** 703 (1902).

J. Tafel, Ann. d. Physik **11,** 613 (1903).

L. Houllevigue, Jour. de Phys. **4 S,** 2 (1903).

G. C. Schmidt, Ann. d. Physik **13,** 622 (1904).

J. Tafel, Ann. d. Physik **14,** 206 (1904).

L. Holborn und L. Austin, Wissensch. Abhandl. d. Phys. -techn. Reichsanstalt, **IV,** 101 (1904).

S. Valentiner und R. Schmidt, Ann. d. Physik **18,** 187 (1905).

J. J. Thomson, Proc. Cambridge Phil. Soc. **XIII,** 212 (1905); Chem. C.-B. **I,** 891 (1906).
Ch. Maurain, C. R. **141,** 1223 (1905).
J. Stark, "Electricity in Gases".
J. J. Thomson, "Conduction through Gases".
W. Ramsey, Proc. Roy. Soc. **77,** 148 (1906), Zeits. f. Physik **55,** 537 (1906).
V. Kohlschutter und R. Muller, Zeits. f. Elektroch. **12,** 365 (1906).
V. Kohlschutter, Zeits. f. Elektroch. **12,** 869 (1906).
L. Houllevigue, Jour. de Phys. **6,** 596 (1907).
V. Kohlschutter und Th. Goldschmidt, Zeits. f. Elektroch. **14,** 221 (1908).
Th. Goldschmidt, Inaug.-Dissert. Strasburg i.E. (1908).
Fr. Fischer und O. Hahnel, Zeits. f. Elektroch. **14,** 366 (1908).
V. Kohlschutter, Zeits. f. Elektroch. **14,** 417 (1908).
Fr. Fischer und O. Hahnel, Zeits. f. Elektroch. **14,** 433 (1908).
V. Kohlschutter, Zeits. f. Elektroch. **14,** 437 (1908).
V. Kohlschutter, Zeits. f. Elektroch. **14,** 681 (1908).
B. Walter, Zeits. f. Electroch. **14,** 695 (1908).
F. Fisher, O. Haehnel Zeits. f. Electroch. **14,** 677 (1908).
Fr. Fischer, Zeits. f. Elektroch. **15,** 512 (1908).
J. Stark, Zeits. f. Elektroch. **14,** 754 (1908).
J. Stark, Zeits. f. Elektroch. **15,** 509 (1909).
J. Stark, Univ. Berlin, Zeits. f. Electroch. **15,** (1909).
Pfund, Phys. Rev. **28,** 324 (1909).
J. Stark, Zeits f. Electroch. **15,** 509 (1909).
V. Kohlschutter, Zeits. f. Electroch. **15,** 316 (1909).
L. Houllevigue, C. R. **148,** 1320 (1909) and 1237 (1910).
V. Kohlschutter, Zeits. f. Elektroch. 930 (1910).
V. Kohlschutter, Zeits. f. Elektroch. **17,** 393 (1911).
H. V. Dechend und W. Hammer, Zeits. f. Elektroch. **17,** 235 (1911). (Ber. d. Naturf. Ges. zu Freiburg i.B. **18,** 127); vgl. auch Jahrb. d. Rad. und Elektr. **8,** 83 (1911).
Kohlschutter Jahrbuch. d. Radioakt **9,** 355 (1912).
H. B. Williams, Phys. Rev. **2,** 402 (1913).
O. Steehlman, Phys. Rev. **6,** 330 (1913).
C. W. C. Kaye, Proc. Phys. Soc. Lon. **25,** 198 (1913).
C. Bergholm, Am. Physik **43,** 1–23 (1914).
W. Planck, Phys. Zeits. 563 (1914).
W. F. G. Swann, Phil. Mag. **28,** 467 (1914).
A. M. Tyndall and H. G. Hughes, Phil. Mag. **27,** 415 (1914).
B. Pogany, Phys. Zeits. **15,** 688 (1914).
Nicholson, Phys. Rev. **3,** 1 (1914).
H. B. Williams, Phys. Rev. **4,** 517 (1914).
B. Pogany, Phys. Zeits. **17,** 251 (1916).
W. Reinders and L. Hamburger, K. Akad. Amst. Proc. **19,** 958 (1917).
R. W. King, Phys. Rev. **10,** 292 (1917).
Fabry and Bursson, Jour. de. Phys. June, (1919).
J. A. Becker, L. F. Curtiss, Phys. Rev. **15,** 457 (1920).
J. J. Thompson, "Rays of Positive Electricity," 121 (1921).
H. Kahler, Phys. Rev. **18,** 210 (1921).
F. Gross, Zeits. f. Physik **6,** 376 (1921).
Kingdon and Langmuir, Phys. Rev. **20,** 107 (1922).
F. Gross, Zeits. f. Physik **16,** 40–5 (1923).
K. H. Kingdon, Phys. Rev. **21,** 210 (1923).
Kingdon and Langmuir, Phys. Rev. **22,** 148 (1923).
Res. Staff G. E. (London) Phil. Mag. **45,** 98 (1923).
K. Lauch, Ann. d. Physik **74,** 55 (1924).
S. S. MacKeown, Phys. Rev. **23,** 85 (1924).
L. R. Ingersoll and S. S. de Vinney, Phys. Rev. **26,** 86 (1925).
J. Mazur, Bull. Acad. Polonaise, 81 (1925).
H. Dewhurst, Phys. Soc. **39,** 39 (1926).
A. Guntherschutze, Zeits. f. Physik **36,** 8, 563 (1926).
L. A. Fourdes, Ind. Assoc. Cult. Sci. Proc. 323 (1926).
A. Guntherschutze, Zeits. f. Physik **38,** 8, 575 (1926).
E. Blechaschmidt, Ann. d. Physik **81,** 999 (1926).
A. v. Hippel, Ann. d. Physik **80,** 7, 672 (1926).
F. Joliot, Compt. Rend. 196 (1926).
"Cathode Disintegration" Engineering 410 (1927).
F. W. Warburton, Phys. Rev. **30,** 673 (1927).
B. Steinberg, Phys. Rev. **29,** 352 (1927).
A. Andant, Rev. d'Optique **6,** 364 (1927).
A. Andant, Rev. d'Optique **6,** 81, 1043 (1927).
T. Baum, Zeit. f. Physik **40,** 9, 686 (1927).
A. Ferry, Jour. de. Phys. **9,** 39 (1928).
A. Ferry, Compt. Rend. 187 (1928).
H. Nagaoka and T. Fretagami, Imp. Acad. Tokyo Proc. **4,** 20 (1928).

L. R. Ingersoll and L. O. Sordahl, Phys. Rev. **32,** 649 (1928).
R. S. Bartlett, Phil. Mag. **5,** 848 (1928).
L. R. Ingersoll and J. D. Hanawalt, Phys. Rev. **34,** 972 (1929).
T. Asada and K. Quasaborth, Zeits. f. physik. Chem. **5–6,** 143 (1929).
Sophie Dembenska, Zeits. f. Physik **54,** 1 and 2 (1929).
F. Smith. U.S.P. **1,** 710. 747 (1929).
A. Knapp, Am. Cer. Soc. Jour. **13,** No. 11 (1930).
L. R. Ingersoll, Nature, **126,** 204 (1930).
L. Van Atta, Rev. Sci. Inst. **1,** 687 (1930).
F. Joliot, Compt. Rend. **190,** 627 (1930).
E. Perucca Ann. d. Physik **4, 2,** 252 (1930).
C. H. Cartwright, Rev. Sci. Inst. **1,** 758 (1930).
F. Joliot, Ann. d. Physik **15,** 418 (1931).
Klemensiewicz and Waswicz, Zeits. f. Physik **71,** 11 and 12, 817 (1931).

ELECTROMAGNETIC EQUATIONS AND SYSTEMS OF UNITS

By Leigh Page
Yale University

(Received March 2, 1932)

Abstract

The electromagnetic equations are put in their most general form and the following criteria for a useful set of equations and sets of units are established: (1) there should be only one set of fundamental equations, (2) there should be two sets of units, one c .g. s. for the physicist and one practical for the engineer, (3) the eight internationally recognized units should be retained in the practical system, (4) the fundamental equations should be symmetrical in form. It is shown that these criteria can be fulfilled in an infinite variety of ways, of which the Heaviside-Lorentz system of equations provides the simplest. The relations between c. g. s. and practical units employing this system of equations are tabulated.

THE subject of electromagnetic units has received considerable attention in recent years, and proposals for new and more convenient sets of units have been made on a number of occasions.[1] In particular this matter has been discussed in some detail by the International Electrotechnical Commission, which is reported[2] to have adopted certain hypotheses which the writer believes will only complicate an already confused situation. Many of the proposals made seem to be concerned only with the units themselves and to ignore the far more vital question of the *form* of the fundamental equations of electromagnetic theory. Again, many of the discussions become submerged in the question of the physical dimensions of electric and magnetic quantities, and a feeling is evident that other dimensions than those of mass, distance and time must be involved, particularly in connection with the dielectric constant κ and the permeability μ. In this connection it does not seem to be recognized that the introduction of physical dimensions is purely a matter of convenience, and, as is shown by Planck[3] and again by Bridgman,[4] that it is quite possible to eliminate all physical dimensions from physical equations. The purpose of this communication is (1) to point out that the introduction of κ and μ into the equations of electromagnetic theory does not necessarily imply any new physical dimensions, (2) to investigate the most general form in which the fundamental equations may be written, (3) to formulate and attempt to justify the criteria which a useful set of equations and sets of units should satisfy, and (4) to show how these criteria may be fulfilled.

[1] See Bibliography in "Historical Outline of the Electrical Units", by A. E. Kennelly, Jour. Eng. Educ. **19**, Nov. (1928).

[2] Magnetic-Circuit Units as Adopted by the I. E. C., by A. E. Kennelly, Winter Convention of A.I.E.E. (1931).

[3] Planck's Heat Radiation, trans. by Masius, p. 174 (1914).

[4] Dimensional Analysis, P. W. Bridgman, Chap. VIII (1922).

In the first place it is essential to emphasize the fact that the necessity of distinguishing two electrical vectors ($\boldsymbol{D}$ and $\boldsymbol{E}$) and two magnetic vectors ($\boldsymbol{B}$ and $\boldsymbol{H}$) is due to the existence of polarizable media —dielectrics in the electric case and permeable media in the magnetic case. Therefore the logical method of approach is to write down the field equations for charges and currents in empty space and then deduce the equations for material media from our knowledge of the structure of matter.

Unfortunately we are accustomed to employ different systems of units, and worse yet different fundamental equations, according as our interest is centered in electrical or in magnetic quantities. To avoid confusion, let us denote e.s.u. by the subscript s and e.m.u. by the subscript m. Then, in developing the theory of dielectrics we find it convenient to introduce a vector[5] $\boldsymbol{D}_s'$ *defined* in terms of the electric intensity $\boldsymbol{E}_s$ and the polarization $\boldsymbol{P}_s$ by the relation

$$\boldsymbol{D}_s' = \boldsymbol{E}_s + 4\pi\boldsymbol{P}_s. \tag{1}$$

As this equation must be physically homogeneous, it follows that $\boldsymbol{D}_s'$, $\boldsymbol{E}_s$ and $\boldsymbol{P}_s$ are all of the *same* physical dimensions.

Now in isotropic media, $\boldsymbol{D}_s'$ is proportional to $\boldsymbol{E}_s$ and has the same direction. Consequently we may write

$$\boldsymbol{D}_s' = \kappa_s\boldsymbol{E}_s. \tag{2}$$

The dielectric constant or permittivity so defined is necessarily a pure number which has the value unity in empty space where $\boldsymbol{P}_s$ vanishes. It must be remembered, though, that (2) is a special relation valid only in isotropic media. It does not hold in crystalline media, for instance. On the other hand (1) is a general relation of universal validity.

In e.m.u. (1) and (2) become

$$\boldsymbol{D}_m' = \frac{1}{c^2}\boldsymbol{E}_m + 4\pi\boldsymbol{P}_m, \tag{3}$$

$$\boldsymbol{D}_m' = \frac{\kappa_s}{c^2}\boldsymbol{E}_m, \tag{4}$$

where c (the ratio of the units of charge) is $2.99796(10)^{10}$ cm/sec. Now *either* Eq. (4) may be left as it is, *or* we may define a permittivity κ_m in e.m.u. by

$$\boldsymbol{D}_m' = \kappa_m\boldsymbol{E}_m, \quad \kappa_m = \frac{\kappa_s}{c^2}.$$

Evidently κ_m has the physical dimensions $[T^2/L^2]$. It is purely a matter of convenience as to whether we adopt the first or the second of these proce-

[5] The D_s' here used differs from the conventional definition of electric displacement by a numerical factor. It is, however, much to be preferred on account of the exact analogy it affords between the electric and magnetic cases.

dures. Since, however, the permittivity is usually tabulated in e.s.u. and since κ_s is a pure number, the first is clearly preferable.

Similar considerations apply in the magnetic case. If we adopt the hypothesis that the properties of permeable media are due to the existence of actual magnetic dipoles (a point of view long discarded by the physicist and one quite incapable of explaining such phenomena as diamagnetism, magnetization by rotation, etc.) the magnetic intensity $\boldsymbol{H}_m$ is the fundamental vector appearing in the field equations for empty space, and we find it convenient in discussing the properties of permeable media to introduce a vector $\boldsymbol{B}_m$ *defined* by the equation

$$\boldsymbol{B}_m = \boldsymbol{H}_m + 4\pi \boldsymbol{I}_m, \tag{5}$$

where $\boldsymbol{I}_m$ is the intensity of magnetization. On the other hand, if we adopt the universally accepted view that magnetic properties are due to the presence of Ampèrian circuits in the atoms or molecules of magnetic substances, then $\boldsymbol{B}_m$ is the fundamental vector appearing in the field equations and (5) *defines* $\boldsymbol{H}_m$. It is clear from (5) that $\boldsymbol{B}_m$, $\boldsymbol{H}_m$ and $\boldsymbol{I}_m$ have the same physical dimensions. In the case of an isotropic paramagnetic or diamagnetic medium $\boldsymbol{B}_m$ is proportional to $\boldsymbol{H}_m$, so we may write

$$\boldsymbol{B}_m = \mu_m \boldsymbol{H}_m \tag{6}$$

to define the permeability μ_m, which is a pure number. A relation of this type can be applied to ferromagnetic media only as a rough approximation, since in such media $\boldsymbol{B}_m$ is not even a one-valued function of $\boldsymbol{H}_m$. While (6) is a special relation of only limited validity, (5) is a general relation which always holds.

Now, if we write (5) and (6) in e.s.u. we get

$$\boldsymbol{B}_s = \frac{1}{c^2}\boldsymbol{H}_s + 4\pi \boldsymbol{I}_s, \tag{7}$$

$$\boldsymbol{B}_s = \frac{\mu_m}{c^2}\boldsymbol{H}_s. \tag{8}$$

Again we are at liberty *either* to content ourselves with (8) *or* to define a permeability μ_s in e.s.u. by

$$\boldsymbol{B}_s = \mu_s \boldsymbol{H}_s, \quad \mu_s = \frac{\mu_m}{c^2},$$

where μ_s has the dimensions $[T^2/L^2]$. Evidently the first alternative is to be preferred for the same reasons as in the case of the permittivity.

The purpose of the foregoing discussion is to show the meaninglessness of the dogmatic assertion[2] that permeability has or has not physical dimensions. A far more important question and one which is necessarily involved in the determination of the physical dimensions of κ and μ is the choice of a suitable set of fundamental equations and of convenient sets of units to employ in them. The present situation is deplorable in that the *four* systems of

units commonly employed demand *three* distinct sets of equations. Thus the force between two point charges *in vacuo* is given by

$$F = \frac{q_e q_e'}{r^2} \qquad \text{in e.s.u.,}$$

$$F = c^2 \frac{q_m q_m'}{r^2} \text{ in e.m.u. or } Q.E.S. \text{ (practical units),}$$

$$F = \frac{q_l q_l}{4\pi r^2} \text{ in h.l.u. (Heaviside-Lorentz units).}$$

This situation would be immensely aggravated if the engineering profession adopted the Giorgi or Dellinger-Bennett units, for then the student of electricity would have to memorize *four* distinct sets of equations instead of *three* as at present.

If we write the fundamental equations of electromagnetic theory in terms of the electromagnetic units employed by Maxwell, we have

$$\nabla \cdot \boldsymbol{D}_m = \rho_m, \tag{9}$$

$$\nabla \cdot \boldsymbol{B}_m = 0, \tag{10}$$

$$\nabla \times \boldsymbol{H}_m = 4\pi(\dot{\boldsymbol{D}}_m + \rho_m \boldsymbol{v}), \tag{11}$$

$$\nabla \times \boldsymbol{E}_m = -\dot{\boldsymbol{B}}_m, \tag{12}$$

$$\boldsymbol{D}_m = \frac{\boldsymbol{E}_m}{4\pi c^2} + \boldsymbol{P}_m = \frac{\kappa_s}{4\pi c^2}\boldsymbol{E}_m, \tag{13}$$

$$\boldsymbol{B}_m = \boldsymbol{H}_m + 4\pi \boldsymbol{I}_m = \mu_m \boldsymbol{H}_m, \tag{14}$$

$$\boldsymbol{F} = \rho_m \{\boldsymbol{E}_m + \boldsymbol{v} \times \boldsymbol{B}_m\}, \tag{15}$$

where ρ_m is electric charge density and $\boldsymbol{F}$ represents force per unit volume. The κ_s and μ_m appearing in these equations are the same dimensionless ratios used previously herein, having the value unity in empty space.

We can change the form of these equations by replacing each of the electromagnetic quantities involved by a new quantity multiplied by an arbitrary numerical or dimensional constant. These new quantities will be denoted by letters without subscripts. This most general transformation, however, would be profitless because it would alter certain definitions which are universally considered desirable to retain. Thus $\boldsymbol{E}$ is defined as force per unit charge and $\boldsymbol{H}$ as force per unit pole, $\boldsymbol{P}$ is electric moment per unit volume and $\boldsymbol{I}$ magnetic moment per unit volume. So the most general transformation to be considered must be subject to the relations

$$\rho \boldsymbol{E} = \rho_m \boldsymbol{E}_m, \qquad \sigma \boldsymbol{H} = \sigma_m \boldsymbol{H}_m,$$

$$\frac{\boldsymbol{P}}{\rho} = \frac{\boldsymbol{P}_m}{\rho_m}, \qquad \frac{\boldsymbol{I}}{\sigma} = \frac{\boldsymbol{I}_m}{\sigma_m},$$

where σ is the fictitious magnetic charge density, $-\nabla \cdot \boldsymbol{I}$. For the first pair represents a force per unit volume and the second a length.

We are at liberty, then, to introduce four arbitrary constants which may be dimensional. Without loss of generality we may put

$$\rho_m = \frac{\alpha}{c}\rho, \quad \sigma_m = \beta\sigma, \quad \boldsymbol{D}_m = \frac{\gamma}{4\pi c}\boldsymbol{D}, \quad \boldsymbol{B}_m = \delta\boldsymbol{B},$$

the particular choice of factors being designed to make the equations symmetrical in c and 4π. Then the fundamental equations become

$$\nabla \cdot \boldsymbol{D} = 4\pi\frac{\alpha}{\gamma}\rho, \tag{16}$$

$$\nabla \cdot \boldsymbol{B} = 0, \tag{17}$$

$$\nabla \times \boldsymbol{H} = \frac{1}{c}\beta\gamma\dot{\boldsymbol{D}} + \frac{4\pi}{c}\alpha\beta\rho\boldsymbol{v}, \tag{18}$$

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\alpha\delta\,\dot{\boldsymbol{B}}, \tag{19}$$

$$\boldsymbol{D} = \frac{1}{\alpha\gamma}\boldsymbol{E} + 4\pi\frac{\alpha}{\gamma}\boldsymbol{P} = \frac{\kappa_s}{\alpha\gamma}\boldsymbol{E}, \tag{20}$$

$$\boldsymbol{B} = \frac{1}{\beta\delta}\boldsymbol{H} + 4\pi\frac{\beta}{\delta}\boldsymbol{I} = \frac{\mu_m}{\beta\delta}\boldsymbol{H}, \tag{21}$$

$$\boldsymbol{F} = \left\{\rho\boldsymbol{E} + \frac{\alpha\delta}{c}\boldsymbol{v} \times \boldsymbol{B}\right\}. \tag{22}$$

From (16), (20) and (22) we see at once that the force between two point charges q_1 and q_2 at rest a distance r apart is

$$F = \alpha^2\frac{q_1 q_2}{\kappa_s r^2}, \tag{23}$$

and that the physical dimensions of q are

$$[q] = \left[\frac{M^{1/2}L^{3/2}}{\alpha T}\right], \tag{24}$$

since κ_s is a pure number. The electrostatic system of equations with the definition of $\boldsymbol{D}$ given in the early part of this paper is obtained by making $\alpha = \gamma = 1$, $\beta = \delta = c$. On the other hand e. s. u. with Maxwell's definition of $\boldsymbol{D}$ require $\alpha = 1$, $\gamma = 4\pi$, $\beta = \delta = c$.

With this general formulation of the fundamental equations of electromagnetic theory at our disposal, the next step is to state the criteria which should govern our choice of the arbitrary constants α, β, γ, δ and determine our selection of sets of units. It seems to the author that the desirability of the following criteria is almost self-evident.

(1) There should be one and only one set of fundamental equations in general use by both physicists and engineers. This simplification would do away with all necessity of converting from the electrostatic to the electromagnetic or Heaviside-Lorentz system or vice-versa, and would abolish at one stroke the endless confusion caused the student in having to memorize three sets of equations.

(2) There should be two sets of units, one c. g. s. for the physicist and the other a practical set for the engineer. Both sets, however, should employ the *same* set of fundamental equations, as is the case with the e. m. u. and Q. E. S. systems at present. The reasons for this criterion are so obvious as to require no comment.

(3) The eight internationally recognized practical units (joule, watt, coulomb, ampere, volt, ohm, farad, henry) must be retained in the proposed practical system of units. This requirement is necessitated by the long established use of these units and the fact that electrical instruments are generally calibrated in terms of them.

(4) The set of equations chosen should exhibit the symmetry existing between electrical quantities on the one hand and magnetic quantities on the other. The importance of this criterion is sometimes overlooked. To the worker who is using the entire set of electromagnetic equations, however, it is almost vital. No more complicated theory than that involved in the deduction of the wave equation or of the expression for radiation pressure is needed to convince oneself of the vast superiority of a symmetrical set of equations. From the practical point of view a great simplification is effected as well, for a symmetrical set of equations halves the number of different coefficients appearing in the equations and halves the number of different relations between practical and c. g. s. units.

Our next step is to investigate the means by which these criteria can be satisfied. It will be seen that they can be fulfilled in a variety of ways, but reasons will be advanced for favoring one solution over all the others. We can classify possible transformations as (1) those for which the arbitrary constants $\alpha, \beta, \gamma, \delta$ are pure numbers and (2) those for which these constants or some of them are dimensional. All transformations of the first class leave unaltered the physical dimensions of the electromagnetic quantities, whereas those of the second class introduce changes in dimensions.

It is, however, more convenient to classify possible transformations with regard to the degree of symmetry which they produce in the fundamental equations. We shall designate as *partial symmetry* that which makes all terms in (18) and (20) symmetrical with corresponding terms in (19) and (21) respectively. This leaves undetermined the coefficient of the second term on the right of (18) to which there is no corresponding term in (19). Partial symmetry, then, demands that $\beta\gamma=\alpha\delta$, $\alpha\gamma=\beta\delta$, $\alpha/\gamma=\beta/\delta$; that is, $\beta=\alpha$, $\delta=\gamma$. The most important set of equations of this type is that proposed by Gauss, for which $\alpha=\gamma=1$. It is to be noted that Gauss' system retains the electrostatic unit of charge and the electromagnetic unit of pole strength.

Its chief defect is that $\dot{\boldsymbol{D}}$ and $\rho\boldsymbol{v}$ in Ampère's law (18) do not have the same coefficient.

Another interesting set of equations exhibiting partial symmetry is that which makes the coefficients of $\dot{\boldsymbol{D}}$ in Ampère's law (18) and $\dot{\boldsymbol{B}}$ in Faraday's law (19) equal to unity and makes the coefficient of the current density $\rho\boldsymbol{v}$ in (18) equal to 4π as in e. m. u. These specifications require $\alpha=\gamma=(c)^{1/2}$. If we define the permittivity as the ratio of D to E and the permeability as that of B to H, this system makes the permittivity and the permeability of empty space each equal to $1/c$.

Next we shall investigate the conditions for *complete symmetry*. By this term we imply not only symmetry as regards corresponding terms in (18) and (20) on the one hand and in (19) and (21) on the other, but also symmetry as regards $\dot{\boldsymbol{D}}$ and $\rho\boldsymbol{v}$ in (18). The requirements here are more stringent than in the case of partial symmetry, for we have the additional condition $\gamma=4\pi\alpha$ which is required to make the coefficients of $\dot{\boldsymbol{D}}$ and $\rho\boldsymbol{v}$ the same in (18). Hence $\beta=\alpha$, $\delta=\gamma=4\pi\alpha$, and α is the only parameter which remains arbitrary. The case of complete symmetry has obvious advantages over that of partial symmetry, and as our further discussion will be confined largely to this case we shall rewrite the electromagnetic equations in the form which they take under these conditions:

$$\nabla\cdot\boldsymbol{D}=\rho, \tag{25}$$

$$\nabla\cdot\boldsymbol{B}=0, \tag{26}$$

$$\nabla\times\boldsymbol{H}=\frac{4\pi\alpha^2}{c}(\dot{\boldsymbol{D}}+\rho\boldsymbol{v}), \tag{27}$$

$$\nabla\times\boldsymbol{E}=-\frac{4\pi\alpha^2}{c}\dot{\boldsymbol{B}}, \tag{28}$$

$$\boldsymbol{D}=\frac{1}{4\pi\alpha^2}\boldsymbol{E}+\boldsymbol{P}=\frac{\kappa_s}{4\pi\alpha^2}\boldsymbol{E}, \tag{29}$$

$$\boldsymbol{B}=\frac{1}{4\pi\alpha^2}\boldsymbol{H}+\boldsymbol{I}=\frac{\mu_m}{4\pi\alpha^2}\boldsymbol{H}, \tag{30}$$

$$\boldsymbol{F}=\rho\left\{\boldsymbol{E}+\frac{4\pi\alpha^2}{c}\boldsymbol{v}\times\boldsymbol{B}\right\}. \tag{31}$$

In addition to these equations we have (23), which holds for all systems of equations.

The Heaviside-Lorentz system is one of the simplest set of equations exhibiting complete symmetry. It is in general use today, the majority of papers on electromagnetic theory being written in terms of this set of equations. In this system $\alpha=1/(4\pi)^{1/2}$, so that the factor 4π disappears from every one of the fundamental equations. The ratios of D to E and of B to H become respectively the pure numbers κ_s and μ_m each of which is unity in empty space. The c. g. s. unit of charge, as is evident from (23), is that charge which

repels a like equal charge at a distance of 1 cm *in vacuo* with a force of $1/4\pi$ of a dyne. It is of dimensions $[M^{1/2}L^{3/2}/T]$. Similar statements are true of the unit of pole strength.[6] The reader can easily work out the dimensions of any other of the electromagnetic quantities from the fundamental equations.

A set of completely symmetrical equations which retains the electrostatic unit of charge and the electromagnetic unit of pole strength is obtained by making $\alpha = 1$. This system, however, leaves unnecessary 4π's in the fundamental equations and makes both the ratio of D to E and that of B to H equal to $1/4\pi$ in empty space. As α is a pure number in both the Heaviside-Lorentz system and in this system, the physical dimensions of all quantities are the same in the two.

Kennelly[7] has proposed a system in which $\alpha = (c/4\pi)^{1/2}$. He calls this an equidimensional system, apparently not realizing that all systems of equations of the symmetrical type are equidimensional in the sense in which he employs the word. As α is a dimensional constant, the physical dimensions of the electromagnetic quantities are different in this system from those of the two previous systems discussed. Charge, for instance, is of dimensions $[M^{1/2}L/T^{1/2}]$ and the c. g. s. unit of charge is that charge which repels a like equal charge at a distance of 1 cm in vacuo with a force of $3(10)^{10}/4\pi$ dyne. The coefficients of the terms on the right-hand sides of (27) and (28) become unity, but the ratio of D to E and of B to H are $1/c$ in empty space. Like the Heaviside-Lorentz system this system removes the factor 4π from every one of the fundamental equations. Whereas the former system leaves the c's in (27) and (28) (where they properly belong, since these equations lead directly to the wave equation) and removes them from (29) and (30), Kennelly's system removes the c's from (27) and (28) and inserts them in (29) and (30) (where there is no good reason for their appearance). So while Kennelly's system has many of the advantages of the Heaviside-Lorentz system, the latter has the theoretical advantage of placing the c's where they would naturally be expected and the pragmatic advantage of being a system already widely used. Again it is a great simplification to have the ratio of D to E and that of B to H equal to unity in empty space, and to have the coefficients of $\boldsymbol{E}$ and $\boldsymbol{P}$ the same in (29) and those of $\boldsymbol{H}$ and $\boldsymbol{I}$ the same in (30).

Among the infinitely many completely symmetrical systems of fundamental equations the two simplest are clearly the Heaviside-Lorentz system and the Kennelly system. These two differ only in the equations to which they assign the dimensional constant c, which must appear somewhere in any complete set of equations. The Heaviside-Lorentz system, however, has the several advantages over the Kennelly system cited above.

[6] Occasionally the statement is made that two different physical quantities must not have the same physical dimensions. Are, then, work and moment of force the same physical quantity? The absurdity of this contention is obvious if one considers the possibility of eliminating all physical dimensions, referred to earlier in this paper. Then *all* physical quantities become of the same (i.e., zero) physical dimensions.

[7] A. E. Kennelly, Proc. Nat. Acad. Sci. **17**, p. 147 (1931). The numerical part of α is not made altogether clear in this paper, the author being concerned mainly with the physical dimensions of the quantities involved.

Having examined the various forms into which the fundamental equations may be thrown, the next matter to investigate is the subject of units. Having selected one of the systems of equations described, can we choose practical units of mass, length and time such that the eight internationally recognized practical units of electromagnetic quantities shall be retained unaltered in value? On account of the pairs joule, watt and coulomb, ampere it is clear that we are required to retain the second as the unit of time. Denoting practical units by the subscript p, we may choose new units of mass and length in accord with the relations

$$[M_p] = \bar{\phi}[M] = \bar{\phi}\ \text{gm}, \quad [L_p] = \bar{\theta}[L] = \bar{\theta}\ \text{cm},$$

the bar over a letter signifying the numeric or numerical measure of the quantity whose unit appears in brackets. We have then but two parameters at our disposal. These are sufficient in general, however, to built up a practical system of units with the desired characteristics using *any* system of equations we may choose to select. For if $\bar{\phi}$ and $\bar{\theta}$ are given such values as to make the joule the unit of energy and the coulomb the unit of charge the six remaining units assume automatically the desired values. This follows from the fact that each of the six quantities involved is *defined* in terms of energy U and charge q or in terms of other quantities already defined in terms of these two, plus, of course, time t, the unit of which has been retained unaltered. Thus power $P = U/t$, current $i = q/t$, potential or e.m.f. $V = U/q$, capacitance $C = V/q$, resistance $R = V/i$, inductance $L = V/-(di/dt)$.

The relation between $\bar{\theta}$ and $\bar{\phi}$ necessary to make the joule the unit of energy and the watt the unit of power is the same for all systems of electromagnetic equations, since these are mechanical units. If we denote the numeric in the practical system by $\overline{U}_p$ and that in the c. g. s. system by $\overline{U}$ in the case of energy, then

$$\overline{U}\left[\frac{ML^2}{T^2}\right] = \overline{U}_p\left[\frac{M_pL_p^2}{T_p^2}\right] = \overline{\phi\theta^2}\,\overline{U}_p\left[\frac{ML^2}{T^2}\right],$$

and

$$\overline{U}_p = \frac{1}{\overline{\phi\theta^2}}\,\overline{U}. \tag{32}$$

But, if the joule is to be the practical unit,

$$\overline{U}_p = (10)^{-7}\overline{U}.$$

Consequently

$$\overline{\phi\theta^2} = (10)^7. \tag{33}$$

Next consider the unit of charge. We shall denote the numeric in the practical system by $\bar{q}_p$ and that in the c. g. s. system by $\bar{q}$ and where, as here, it is convenient to distinguish between the numeric and the unit of α, we shall denote the former by $\bar{\alpha}$ and the latter by $[\alpha]$. Then we have

$$\bar{q}\left[\frac{M^{1/2}L^{3/2}}{\alpha T}-\right]=\bar{q}_p\left[\frac{M_p{}^{1/2}L_p{}^{3/2}}{\alpha_p T_p}\right]=\bar{\phi}^{1/2}\bar{\theta}^{3/2}\bar{q}_p\left[\frac{M^{1/2}L^{3/2}}{\alpha_p T}\right],$$

and

$$\bar{q}_p=\frac{1}{\bar{\phi}^{1/2}\bar{\theta}^{3/2}}\left[\frac{\alpha_p}{\alpha}\right]\bar{q}.$$

Since we wish to retain the coulomb as the practical unit of charge,

$$\bar{q}_p=\frac{1}{2.99796(10)^9}\,\bar{q}_s$$

where $\bar{q}_s$ is the numeric in e. s. u. But as the electrostatic system is obtained by making α equal to unity in (23),

$$\bar{q}_s=\bar{\alpha}\bar{q}.$$

Eliminating the $\bar{q}$'s from the last three equations

$$\bar{\phi}^{1/2}\bar{\theta}^{3/2}=\frac{\bar{\eta}}{\bar{\alpha}}(10)^9\left[\frac{\alpha_p}{\alpha}\right] \tag{34}$$

where $\bar{\eta}$ has been put for the number 2.99796. Since (34) depends only on (23), which was obtained from the most general form of the fundamental equations, a system of practical units which satisfies (32) and (34) will contain the eight internationally recognized practical units no matter what system of fundamental equations are selected and no matter whether these equations are symmetrical or not.

Nothing is gained by solving (33) and (34) for $\bar{\phi}$ and $\bar{\theta}$ at the present stage, for the ratio $[\alpha_p/\alpha]$ contains $\bar{\phi}$ and $\bar{\theta}$ when α is a dimensional constant. So we must treat separately those cases in which α has different physical dimensions.

The most important systems of equations which have been discussed are those for which α is a pure number. In such cases the ratio $[\alpha_p/\alpha]$ is the number 1 and

$$\left.\begin{aligned}\bar{\phi}&=\left(\frac{\bar{\alpha}}{\bar{\eta}}\right)^4(10)^{-15},\\ \bar{\theta}&=\left(\frac{\bar{\eta}}{\bar{\alpha}}\right)^2(10)^{11}.\end{aligned}\right\} \tag{35}$$

Gauss' system of equations is of this type, $\bar{\alpha}$ having the value unity. Hence a set of practical units employing Gauss' system which includes the eight internationally recognized practical units is obtained by taking as the practical unit of mass $(1/\bar{\eta})^4\,(10)^{-15}$ gm and as the practical unit of length $(\bar{\eta})^2(10)^{11}$ cm. Any system which retains the electrostatic unit of charge as the c. g. s. unit requires the same practical units of mass and length.

In the Heaviside-Lorentz system of equations $\alpha=1/(4\pi)^{1/2}$. So in order to retain the eight internationally recognized practical units we must take

$(1/(4\pi)^{1/2}\bar{\eta})^4\ (10)^{-15}$ gm as the practical unit of mass and $((4\pi)^{1/2}\bar{\eta})^2\ (10)^{11}$ cm as the practical unit of length. As $(4\pi)^{1/2}\bar{\eta} = 10.6275$ this means roughly a unit of mass of $(10)^{-19}$ gm and a unit of length of $(10)^{13}$ cm.

In Kennelly's system $\alpha = (c/4\pi)^{1/2}$. Hence $\bar{\alpha}^2 = (\bar{\eta}/4\pi)\ (10)^{10}$ and

$$[\alpha_p{}^2] = \left[\frac{L_p}{T_p}\right] = \bar{\theta}\left[\frac{L}{T}\right] = \bar{\theta}[\alpha^2]$$

whence

$$\left[\frac{\alpha_p}{\alpha}\right] = \bar{\theta}^{1/2}.$$

This makes (34) become

$$\bar{\phi}\bar{\theta}^2 = 4\pi\bar{\eta}(10)^8. \tag{36}$$

Comparing with (33) we see that we have been led to an inconsistency. In this exceptional case it is impossible to choose new units of mass and length of such magnitudes as to retain the eight internationally recognized units in the practical system. However the form of (36) indicates that a change in the numeric $\bar{\alpha}$ will enable us to make (36) identical with (33). This will result in only a *single* relation between $\bar{\phi}$ and $\bar{\theta}$ and will allow us to choose the practical unit of *either* mass *or* length at will. Leaving $\bar{\alpha}$ undetermined, then, Eq. (36) is replaced by

$$\bar{\phi}\bar{\theta}^2 = \frac{\bar{\eta}^2}{\bar{\alpha}^2}\,(10)^{18} \tag{37}$$

which becomes identical with (33) if $\bar{\alpha} = \bar{\eta}(10)^{11/2}$. This system we shall designate as the *modified Kennelly system*. Since the unit of length, which is employed in calculating flux, is more important than the unit of mass in electromagnetic problems, it is clearly desirable to retain the centimeter as the unit of length in the practical system. As this requires that θ should be unity, the practical unit of mass becomes $(10)^7$ gm and that of force $(10)^7$ dyne. These fundamental units have already been employed in the Dellinger-Bennett system.

In the modified Kennelly system $4\pi\alpha^2 = 40\pi\bar{\eta}c$. Consequently the completely symmetrical system of fundamental equations (25) to (31) acquires awkward numerical or dimensional coefficients in nearly every term. This unfortunate state of affairs can be alleviated by employing the partially symmetrical system of equations, as γ is then arbitrary. Probably the simplest system of equations is obtained by making $\alpha\gamma = 1$. Then

$$\nabla \cdot \boldsymbol{D} = 40\pi\bar{\eta}c\rho, \tag{38}$$

$$\nabla \cdot \boldsymbol{B} = 0, \tag{39}$$

$$\nabla \times \boldsymbol{H} = \frac{1}{c}\{\dot{\boldsymbol{D}} + 40\pi\bar{\eta}c\rho\boldsymbol{v}\}, \tag{40}$$

$$\nabla \times \boldsymbol{E} = -\frac{1}{c}\dot{\boldsymbol{B}}, \tag{41}$$

$$\boldsymbol{D} = \boldsymbol{E} + 40\pi\bar{\eta}c\boldsymbol{P} = \kappa_s \boldsymbol{E}, \tag{42}$$

$$\boldsymbol{B} = \boldsymbol{H} + 40\pi\bar{\eta}c\boldsymbol{I} = \mu_m \boldsymbol{H}, \tag{43}$$

$$\boldsymbol{F} = \rho\left\{\boldsymbol{E} + \frac{1}{c}\boldsymbol{v} \times \boldsymbol{B}\right\}. \tag{44}$$

These equations have the virtue of making the ratios of D to E and of B to H unity in vacuo and of keeping the c's in the circuital equations where they naturally belong. Another possibility consists in making $\alpha\gamma = c$. This procedure removes the c's from (40) and (41) and inserts them in (42) and (43), making the ratios of D to E and of B to H equal to $1/c$ in empty space. This choice of γ seems less desirable than the previous one.

At this point it is advantageous to summarize the results obtained so far. As regards the choice of fundamental equations we have found that there is a single infinity of completely symmetrical systems of equations. The two systems which make the greatest number of coefficients unity are the Heaviside-Lorentz system with $\alpha = (1/4\pi)^{1/2}$ and the Kennelly system with $\alpha = (c/4\pi)^{1/2}$. As regards the simplicity of form of the fundamental equations these two systems are distinctly preferable to any others, and neither one has any essential advantage over the other.

Our investigation of systems of units has shown that in addition to a set of c. g. s. units a set of practical units including the eight internationally recognized units can be used with any system of fundamental equations by choosing appropriate new units of mass and length. The sole exception is the Kennelly system of equations, where the choice of new units of mass and length cannot be made in such a way as to include the internationally recognized units in the practical set. So the Kennelly system of equations is not suited to our purpose. On the other hand this difficulty is obviated in the modified Kennelly system of equations, which has the unique advantage of permitting us to choose arbitrarily either the unit of mass or the unit of length to be employed in the practical system of units. Unfortunately, however, this system necessitates the introduction of awkward coefficients into the fundamental equations, even if we content ourselves with the partially symmetrical group.

These considerations show that the Kennelly system fails to satisfy our criteria, and that the modified Kennelly system, in spite of its advantageous property cited above, leads to a far more complicated system of fundamental equations than the Heaviside-Lorentz system. Hence the Heaviside-Lorentz system of fundamental equations is clearly indicated as the one most deserving of general adoption. This system makes the coefficients of the right-hand sides of (27) and (28) equal to $1/c$, all the other coefficients in the field Eqs. (25) to (30) being unity.

In the accompanying table is given for all the more important electromagnetic quantities the dimensions required by the Heaviside-Lorentz sys-

tem of equations and the ratio of the numeric $\bar{x}_p$ measuring a quantity in practical units to the numeric $\bar{x}$ measuring the same quantity in c. g. s. units.[8] The practical unit is specified in cases where it has received a commonly accepted name. The ratio of the units is the reciprocal of the ratio of the numerics. The symbol ϵ in the table stands for the number

$$\epsilon = 0.299796(4\pi)^{1/2} = 1.06275.$$

The velocity of light in the practical units under discussion has the numerical value

$$\bar{c}_p = \epsilon^{-2}(10)^{-13}\bar{c} = 0.00265438 = 1/376.735.$$

The permittivity κ_s and the permeability μ_m being pure numbers, have the same values in the practical as in the c. g. s. units. Each is unity in empty space.

The reader's attention is called to the following facts displayed in the table. (1) Corresponding electrical and magnetic quantities have the same physicial dimensions. (2) The conversion ratio $\bar{x}_p/\bar{x}$ for changing from c. g. s. to practical units is the same for corresponding electrical and magnetic quantities. (3) The small number of conversion ratios involved. These properties are characteristic of symmetrical systems of equations.

TABLE I. *Heaviside-Lorentz system of equations.*

Quantity	*Dimensions*	*Practical Unit*	*Ratio $\bar{x}_p/\bar{x}$ of Numerics*
Time	$[T]$	sec	1
Mass	$[M]$		$\epsilon^4(10)^{19}$
Length	$[L]$		$\epsilon^{-2}(10)^{-13}$
Force	$[ML/T^2]$		$\epsilon^2(10)^6$
Energy	$[ML^2/T^2]$	joule	$(10)^{-7}$
Power	$[ML^2/T^3]$	watt	
Charge	$[M^{1/2}L^{3/2}/T]$	coulomb	
Current	$[M^{1/2}L^{3/2}/T^2]$	ampere	$\epsilon^{-1}(10)^{-10}$
Electric flux	$[M^{1/2}L^{3/2}/T]$		
Elect. pot. or e.m.f.	$[M^{1/2}L^{1/2}/T]$	volt	$\epsilon(10)^3$
Electric intensity	$[M^{1/2}/L^{1/2}T]$		
Electric displacement	$[M^{1/2}/L^{1/2}T]$		$\epsilon^3(10)^{16}$
Polarization	$[M^{1/2}/L^{1/2}T]$		
Capacitance	$[L]$	farad	$\epsilon^{-2}(10)^{-13}$
Resistance	$[T/L]$	ohm	$\epsilon^2(10)^{13}$
Pole strength	$[M^{1/2}L^{3/2}/T]$		$\epsilon^{-1}(10)^{-10}$
Magnetic flux	$[M^{1/2}L^{3/2}/T^2]$		
Mag. pot. or m.m.f.	$[M^{1/2}L^{1/2}/T]$		$\epsilon(10)^3$
Magnetic intensity	$[M^{1/2}/L^{1/2}T]$		
Magnetic induction	$[M^{1/2}/L^{1/2}T]$		$\epsilon^3(10)^{16}$
Intensity of magnetization	$[M^{1/2}/L^{1/2}T]$		
Inductance	$[T^2/L]$	henry	$\epsilon^2(10)^{-3}$
Reluctance	$[1/L]$		

[8] It is an astonishing fact that most writers tabulate the relations between units instead of the relations between numerics. In changing from e. m. u. to practical units, for instance, the essential information required is not that the ampere is one-tenth of the e. m. u. of current, but that the numeric giving the magnitude of the current in amperes is ten times the numeric specifying the current in e. m. u.

Since ϵ differs but slightly from unity, this factor can be neglected in rough estimates requiring a transformation from c. g. s. to practical units or *vice versa*. In fact if it were possible to make a change of a few percent in the values of the practical electrical units (excluding the mechanical units joule and watt), this factor could be made rigorously equal to unity.

We shall consider now some of the objections which may be raised to the proposed system of equations and set of practical units. The most serious of these is that some of the fundamental equations as well as the ratios of c. g. s. to practical units demand a knowledge of the velocity of light. This is, however, a characteristic of any *complete* system of equations, whether the electromagnetic system or any other. Furthermore, in view of Michelson's extremely accurate determination of the velocity of light, this objection is not likely to lead to changes in the constants comparable with the errors of even the most precise electrical measurements for a great many years to come.

Next Ampère's law and Faraday's law in the Heaviside-Lorentz system of equations involve the constant factor $(1/c)$. This factor, which is the *only* coefficient other than unity appearing in the equations, is certainly less troublesome than the *two* coefficients 4π and $1/4\pi c^2$ appearing in the electromagnetic system (9) to (15). Finally it is to be noted that the number ϵ does not appear in the equations but is needed only in transforming from c. g. s. to practical units or *vice versa*. So long as computations are conducted entirely in c. g. s. units or entirely in practical units this number is not required. It would be highly desirable to make ϵ unity, but this would require a slight change in the values of the established practical units.

Finally it is appropriate to consider the simplest formulation of electromagnetic theory available if we were subject to neither the restriction imposed by the c. g. s. units nor that imposed by the practical units. Under these circumstances we could attain an ideal degree of simplicity by using either the Heaviside-Lorentz or Kennelly system of equations with new units of length and time so chosen as to make the numerical value of the velocity of light equal to unity. In fact we could, if desired, reduce the number of fundamental units from three to two by giving c the status of a pure number of magnitude unity. This would make the Heaviside-Lorentz and Kennelly systems of equations identical. If we wished to eliminate time and retain only mass and length as fundamental entities, then the unit of time would become a derived unit defined as the time taken by light to travel a unit distance in empty space. Wherever $[T]$ appears in dimensional formulas it would then be replaced by $[L]$. This solution, however, involves too radical a departure from established conventions to be of more than academic interest.

FURTHER STUDIES ON THE THEORY OF SPONTANEOUS DISPERSION OF SMALL LIQUID SYSTEMS WHICH ARE THE SEATS OF PHYSICO-CHEMICAL REACTIONS

By N. Rashevsky
Westinghouse Research Laboratories, East Pittsburgh, Pa.

(Received February 2, 1932)

Abstract

The treatment of the problem which was given in previous publications is generalized so as to include not only the part played by the changes of the surface tension, but also that played by osmotic forces. Whenever any reactions take place in a liquid drop of the kind considered previously, there is always a difference of concentrations inside and outside of the drop. The forces caused by this difference are such that they always tend to divide the drop. The surface tension counteracts this tendency. As a result of this, the drop will divide only above a critical size. Calculations of this critical size leads to dimensions of actual living cells.

IN A series of previous papers[1] we have made a mathematical study of possible causes, which may induce a small growing liquid drop or cell to divide spontaneously in two, after reaching a critical size. The principal factor to which we turned our attention was the variation of the surface tension of the drop with varying size of the latter. We found that this will lead to a spontaneous division under rather general conditions.

The purpose of the present paper is to show that there exists a much more general factor, the influence of which we have hitherto neglected, but which is of a special interest in view of possible biological applications.

If a substance A is soluble in both the drop and the surrounding liquid and if this substance does not undergo any chemical changes neither in the surrounding medium nor in the drop, its concentration c in the drop will be equal to

$$c = \alpha c_0 \tag{1}$$

c_0 being the concentration in the surrounding liquid and

$$\alpha = e^{-\Delta\lambda/RT} \tag{2}$$

where $\Delta\lambda$ is the difference between the heat of solution of A in the drop and in the surrounding liquid. R is the gas constant, T – the absolute temperature.

The distribution (1) of A between the drop and the surrounding solution corresponds to the minimum of the free energy of the system "drop +

[1] N. Rashevsky, Some Theoretical Aspects of the Biological Applications of Physics of Disperse Systems. Physics **1**, 143 (1931). That paper contains also references to the other publications of the author.

surrounding medium". In our previous discussions we have for simplicity neglected the difference of the heats of solution and put $\alpha = 1$.

If however A disappears inside of the drop, being transformed into another substance, then the concentration c of A in the drop is smaller than αc_0, the difference $\alpha c_0 - c$ increasing with the radius of the drop. A division of the drop in two smaller ones therefore is always associated with a *decrease* of $\alpha c_0 - c$, and therefore, since $\alpha c_0 - c = 0$ corresponds to the minimum of the free energy, such a division is always accompanied by a *decrease of the free energy of the system "drop + surrounding liquid"*. The same holds true for the case that A is produced inside of the drop and diffuses into the surrounding liquid. We then have always $c > \alpha c_0$ and the difference $c - \alpha c_0$ increases with the size of the drop. We thus see that the presence of any reaction inside of the drop tends to disperse the drop into smaller ones. The forces which cause this dispersion are due to the difference of osmotic pressures inside and out-

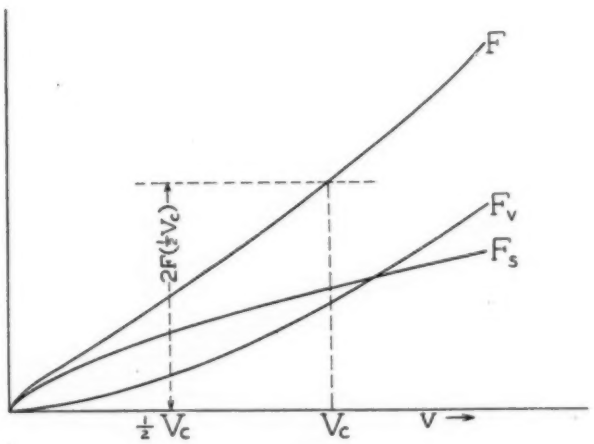

Fig. 1.

side the drop, produced by the difference of c and αc_0. The actual distribution of those forces and their mode of action will vary from case to case, depending on the distribution of A inside the drop.

If besides those no other forces were present the drop would disperse indefinitely and could not exist at all. But the increase of the surface energy which accompanies such a disperson, counteracts it, and therefore as we shall presently see, the division will occur only above a critical size.

If we plot the excess F_v of the free energy of the system over its minimum value, which corresponds to $c = \alpha c_0$, against the volume V of the drop, we shall obtain a curve which is convex toward the v axis. (Fig. 1.) For as we have just seen, the division of the drop results always in a decrease of F_v; hence always

$$F_v(V_1 + V_2) > F_v(V_1) + F_v(V_2). \tag{3}$$

On the other hand assuming for the present a constant surface tension γ, the

free energy of the surface of the drop, F_s, will vary like $V^{2/3}$ (Fig. 1). So that always

$$F_s(V_1 + V_2) < F_s(V_1) + F_s(V_2). \tag{4}$$

The total free energy F is given by the sum of F_v and F_s (Fig. 1) The $F-V$ curve has an inflection point and we seen that for $V<V_c$

$$F(V_1 + V_2) < F(V_1) + F(V_2) \tag{5}$$

but for $V>V_c$

$$F(V_1 + V_2) > F(V_1) + F(V_2). \tag{6}$$

In other words: below V_c the division is followed by an increase of the free energy and cannot occur spontaneously; above V_c- is followed by a decrease of the free energy and will therefore happen spontaneously.

It must be remarked that, strictly speaking, we must include in the total free energy a term expressing the free energy of the substance of the drop itself. This term however, being linear in V, the correction will have no effect on the result.

If γ is not constant, but varies in one of the ways, discussed in reference 1, then the surface forces may also contribute to the division of the drop. But as we see, even in the absence of any variation of γ, *the drop is bound to divide, after reaching a critical size, due to the reactions, which take place in it.*

We shall calculate the size at which the drop will divide, with the assumption that A is formed at a constant rate q inside the drop, and putting again for simplicity $\alpha=1$. We have then:[2]

$$c = c_0 + \frac{qr_0}{3h} + \frac{q(r_0^2 - r^2)}{6D} \tag{7}$$

the following notations being used: r_0—radius of the drop; h—permeability of the drop surface; D—coefficient of diffusion of A inside the drop.

If the concentration of A in an element of volume of the drop dV is c, then dV contains $x=cdV/M$ gr—mols of A, M being the molecular weight of A. A similar element of volume dV in the surrounding medium contains c_0dV/M gr—mols of A. The transfer of dx gr—mols of A from an element of volume in the cell into the surrounding medium liberates

$$RT \log \frac{c}{c_0} dx \text{ ergs} \tag{8}$$

under the assumption of laws of dilute solutions for A.

Hence the excess of free energy in an element of volume dV with a concentration c, over the minimum which corresponds to $c=c_0$ is equal to:

$$RT \int_{c_0 dV/M}^{c\,dV/M} \log \frac{c}{c_0} dx = \frac{RTc_0 dV}{M} \left(\frac{c}{c_0} \log \frac{c}{c_0} - \frac{c}{c_0} + 1 \right) \tag{9}$$

[2] N. Rashevsky, Zur Theorie der spontanen Teilung von mikroskopischen Tropfen II. Zeits. f. Physik **48,** 513 (1928). Also: Physical Aspects of Cellular Growth and Multiplication. Protoplasma **14,** 99 (1931).

remembering that

$$c = \frac{xM}{dV}.$$

In order to obtain the excess F_v of the free energy of the whole drop we must integrate (9) over the whole volume of the latter.

From (7) we have:

$$\frac{c}{c_0} = 1 + \frac{qr_0}{3hc_0} + \frac{q(r_0^2 - r^2)}{6Dc_0} \tag{10}$$

Putting

$$1 + \frac{qr_0}{3hc_0} + \frac{qr_0^2}{6Dc_0} = a \tag{11}$$

and

$$\frac{q}{6Dc_0} = b \tag{12}$$

we have

$$c/c_0 = a - br^2. \tag{13}$$

Introducing this into (9) and remembering that

$$dV = 4\pi r^2 dr \tag{14}$$

we find for the total excess F_v of free energy:

$$F_v = \frac{4\pi RTc_0}{M} \int_0^{r_0} r^2 [(a - br^2) \log (a - br^2) - (a - 1) + br^2] dr. \tag{15}$$

After somewhat elaborate but elementary calculations, in which use is made of the following integrals:

$$\int x^m \log (a + bx) dx = \frac{x^{m+1}}{m+1} \log (a + bx) - \frac{b}{m+1} \int \frac{x^{m+1}}{a + bx} dx$$

$$\int \frac{x^4 dx}{a + bx^2} = \frac{x^3}{3b} - \frac{ax}{b^2} + \frac{a^2}{b^2} \int \frac{dx}{a + bx^2}$$

$$\int \frac{x^6 dx}{a + bx^2} = \frac{x^5}{5b} - \frac{ax^3}{3b^2} + \frac{a^2 x}{b^3} - \frac{a^3}{b^3} \int \frac{dx}{a + bx^2}$$

and

$$\int \frac{dx}{a + bx^2} = \frac{1}{2(-ab)^{1/2}} \log \frac{(a)^{1/2} + x(-b)^{1/2}}{(a)^{1/2} - x(-b)^{1/2}} \text{ for } a > 0, b < 0,$$

we obtain:

$$F_v = \frac{4\pi RTc_0}{M}\left[\left(\frac{ar_0^3}{3} - \frac{br_0^5}{5}\right)\log\left(a - br_0^2\right) + \frac{7b}{25}r_0^5 + \frac{15 - 19a}{45}r_0^3 - \frac{4a^2}{15b}r_0 + \frac{2a^3}{15b(ab)^{1/2}}\log\frac{(a)^{1/2} + r_0(b)^{1/2}}{(a)^{1/2} - r_0(b)^{1/2}}\right]. \tag{16}$$

This expression can be considerably simplified in two cases:

1. $a-1<1$; and $h \gg D$, so that we may put $a=1+br_0^2$.

Expanding the last term of (16) we find that all members of power smaller than 7-th in r_0 cancel out, and we obtain

$$F_v = \frac{4\pi RTq^2}{945Mc_0D^2}r_0^2, \tag{17}$$

2. $a-1<1$ and $D \gg h$, so that $a \eqsim 1+qr_0/3hc_0$, and $b \eqsim 0$.

Then we find, by expanding also the first member in the square brackets of (16):

$$F_v = \frac{2RTq^2\pi}{27Mh^2c_0}r_0^5. \tag{18}$$

When a drop with a radius r_0 divides in two halves, each half has a radius $r_1=0,8\,r_0$. Hence the free energy, liberated in the process of division is, using (17):

$$\delta F_v = \frac{8\pi RTq^2}{315Mc_0D^2}(r_0^7 - 2r_1^7) = \frac{2,4RTq^2}{945Mc_0D^2}r_0^7 \tag{19}$$

or, using (18):

$$\delta F_v = \frac{0,8RTq^2\pi}{27Mh^2c_0}r_0^5. \tag{20}$$

The increase of the surface energy in case of a division in two halves is equal to

$$\delta F_s = 2 \times 4\pi r_1^2\gamma - 4\pi r_0^2\gamma = \pi r_0^2\gamma \tag{21}$$

γ being the surface tension.

For small values of r_0 $|\delta F_s| > |\delta F_v|$, for larger ones—the reverse holds true. The critical value r_c, above which a spontaneous division will occur is given by

$$|\delta F_c| = |\delta F_s|.$$

Using (19) we find

$$r_c = \left(\frac{945Mc_0D^2\gamma}{2,4RTq^2}\right)^{1/5}. \tag{22}$$

Using (20) we obtain

$$r_c = 3\left(\frac{Mh^2c_0\gamma}{0{,}8RTq^2}\right)^{1/3}. \tag{23}$$

The same expressions are obtained if the substance diffuses *into* the drop and disappears there at a constant rate q. For all we need to do is to change the sign of q in (7) which leaves the expressions (22) and (23) unchanged.

Assuming q to be of the order of 10^{-4} gr. cm^{-3} min^{-1}, which is approximately the rate of consumption of oxygen in a living cell, assuming further $D \sim 10^{-5}$ $cm^2 \cdot min^{-1}$, a value found for the diffusion of oxygen through tissues,[3] and taking $\gamma \sim 1$ dyn $\cdot cm^4$ and $c_0 = 3 \times 10^{-4}$ gr $\cdot cm^{-3}$, which is approximately the concentration of oxygen in the atmosphere, we find, using (23), that r_c comes out to be of the order of magnitude of 4×10^{-3} cm, which is the order of magnitude of some living cells. Eq. (17) is justified, because $a - 1 = 1/2$, and the permeability h of the cell for oxygen is very great.[5]

A discussion of the biological aspects of this study will be given elsewhere.

[3] A. Krogh, The Rate of Diffusion of Gases through Animal Tissues. J. of Physiol. **52**, 391 (1919).

[4] E. Newton Harvey, A Determination of the Tension at the Surface of Eggs of the Annelid Chaetopterus. Biol. Bull. **60**, 67 (1931).

[5] E. Newton Harvey, The Permeability of Cells for Oxygen and its Significance for the Theory of Stimulation. J. Gen. Physiol. **5**, 215 (1922).

EDITOR'S COLUMN

Physics in the Soviet Union

READERS of *Physics* will undoubtedly be interested in the appearance of the new journal *Physikalische Zeitschrift der Sowjetunion*, first copies of which reached America a short time ago. It is to be published monthly commencing with the issue of January 1932, the purpose being to furnish an outlet for the articles of Russian physicists who have hitherto been publishing principally in the German periodicals. The articles will be in German, French, or English depending upon the choice of the author, but not in Russian.

As Professor Joffé states in his introduction to the first issue, the intensely practical and technical aims of the Russian social program naturally will be reflected in the character of the papers published, but works of a more problematical and theoretical nature will also appear.

A good share of the first issue is devoted to recent work on the so-called *Sperrschicht* photoelectric effect,[1] and to the results of the conference on this and related topics which was held in Leningrad last September. The phenomenon to which reference is made may be described roughly as follows. A thin layer of a semi-conducting substance such as Cu_2O, Cu_2S, etc. is placed between two thin metallic electrodes of the same kind of metal (copper or platinum for instance) one of which is then illuminated from a source of light. It is found that if the circuit is not closed between the electrodes a difference of potential will be set up between then, and if the circuit is closed a current will flow in it.[2] Depending upon the exact character of the material forming the electrodes and of the semi-conducting layer the irradiated electrode may become either negatively or positively charged (*Vorderwand* or *Hinterwand* effect).[3]

Kurtschatow and Sinelnikow[4] have performed a series of experiments on a number of cells in an endeavor to elucidate the mechanism of the effect. The thesis is that under the action of the light, electrons are liberated within the layer of the semi-conducting material (Sperrschicht) which are then able to migrate into the metal. The effect is thus intimately related to the "inner" or "volume" photoelectric effect of the material forming the layer, which is responsible also for the change in conductivity of certain substances on irradiation, of which selenium forms a notable example. The differences between the *Vorderwand* and *Hinterwand* effects are traced by these authors to the spectral transmission properties of the layer. If the material forming the layer is opaque to the radiation this will be absorbed in the immediate region of the irradiated electrode and the electrons which are set free will migrate to it; if the material of the later partially transmits the radiation then an appreciable part of the light will traverse the layer and liberate electrons in the neighborhood of the other electrode which may then become the negative pole of the cell. For white light or any non-monochromatic radiation both effects are superposed and the irradiated electrode may become either positively or negatively charged.

In a second paper by Kurtschatow, Sinelnikow, and Borissow,[5] the examination of the relation of this effect to the volume photoelectric effect of the layer substance is pushed a little further. The point in question concerns the threshold frequencies for the two

[1] What the proper Englishing of this term is to be I do not know. The literature on the subject which is known to me is almost exclusively in German. In their article (reference 6) which is in English, Frenkel and Joffé refer to the effect as a "contact photoelectric effect".

[2] B. Lange, Phys. Zeits. **31,** 139 (1930). Science Abstracts A, No. 2646 (1930).

Phys. Zeits. **31,** 964 (1930). Science Abstracts A, No. 1381 (1931).

[3] W. Schottky, Phys. Zeits. **31,** 913 (1930); Zeits. f. Tech. Physik **11,** 458 (1930). Science Abstracts A, No. 1380 (1931).

[4] I. W. Kurtschatow, and C. D. Sinelnikow, Sow. Phys. **1,** 23 (1932).

[5] I. Kurtschatow, C. Sinelnikow, and M. Borissow, Sow. Phys. **1,** 42 (1932).

effects; i.e., the minimum frequencies ν of the light for which they are observed. It is to be expected if the theory we have sketched above is true, that the threshold frequency for the contact photoelectric effect will be far below that for the ordinary surface photoelectric effect. In the first case the energy of the photon, $h\nu$, need be large enough to simply release the electron from its relatively weak binding to the atoms *within* the substance, while in the second it must in addition give the electron sufficient energy to help it out of the surface, an additional process requiring roughly 4 electron volts more energy. But it is still uncertain whether or not an electron which is "free" in the inside of the layer material needs an additional amount of energy in order that it can pass the boundary separating the metal electrode from the semi-conducting layer. If not, then the thresholds for the volume and contact photoelectric effects should be identical, and if so then the threshold frequency for the latter should be higher than for the former, since the threshold frequency for the volume photoelectric effect is presumably the minimum frequency which light must have in order to release electrons within the body of the material. From the available data the author's conclude that the latter is the correct alternative, the threshold wavelengths being respectively 41000A and 14000A, corresponding to energies of about 0.3 and 0.9 electron-volts for a cell in which the layer material was Cu_2O.

The theory of the rectifying action of these photocells and of contacts between surfaces in general is discussed in a paper[6] by Frenkel and Joffé, which is an extension of a paper by Frenkel[7] which appeared recently in the *Physical Review*. The theory is on the lines of the physical picture sketched above, but with a detailed study with the help of Sommerfeld's theory of a metal as an electron gas obeying the Fermi statistical distribution. The rectifying action of a contact between two conducting or semi-conducting substances is due to some asymmetry in the transmission of a current through the contact—the current must be smaller with an *emf* in one sense across the contact than it is with the *emf* in the opposite sense. The cause of this asymmetry may be different for different types of contact: it may be due to the difference in the number of free electrons per cm^3 in the two substances on the two sides of the contact, or it may be due to some slight asymmetry in the transmission of the barrier, or to some other cause. The first of these is possibly the major factor in the case of the photo-cells of which we have been speaking, as we must think of the semi-conducting layer material as containing, in general, but very few "free" electrons compared to the metallic electrodes.

We shall not attempt to give a discussion of the remaining papers of the first issue of this new journal, but shall give the (translated) titles for the information of the reader:

L. Landau, "On the theory of energy transfer by collisions".

J. Frenkel, "On a general method of treating incomplete systems in wave mechanics".

B. Hochberg, "The influence of strong electrical fields on the adsorption of water on the surface of mica".

S. Schurkow, "On the influence of adsorbed surface layers on the strength of small quartz fibers".

A. Joffé and A. Walther, "Breaking strengths of thin glass fibers and mica sheets".

A. Joffé, "On the mechanical strength of thin layers".

B. Kurtschatow and M. Eremejew, "On the electrical properties of mixed crystals of Rochelle salt".

A. Joffé, "On the question of the mechanism of electrical breakdown".

E. HILL

[6] J. Frenkel and A. Joffé, Sow. Phys. **1**, 60 (1932).

[7] J. Frenkel, Phys. Rev. **36**, 1604 (1930).

The Neutron

CONSIDERABLE comment has been aroused within the last few weeks by the pronouncement, through the medium of the daily press, of the experimental verification of the existence of the neutron by Professor J. Chadwick of the Cavendish Laboratory of Cambridge University. In the meantime Professor Chadwick has published a short ac-

count of his experiments[1] so that we can now understand a little more clearly the nature of the evidence.

It is perhaps unwise to make a definite statement as to the initial proposer of the (or perhaps we might better say "of a") neutron as it is the type of idea which is likely to occur to people working on widely separated phases of atomic problems. We may mention early suggestions of Rutherford[2] and of Harkins,[3] and more recent ones of Langer and Rosen[4] and of Pauli.[5]

Perhaps the most important property which is desired of a neutron is the possession of a mass but the lack of an electrical charge. Such a particle could then go many places which are practically denied to electrons and protons; e.g., into the nuclei of atoms. If we can conceive of the neutron as built on the lines of ordinary atoms, in order to avoid the necessity of postulating the existence of a third fundamental electrical entity on a par with electrons and protons, then the simplest structure which we can give to it is that of an unusually close combination of an electron and a proton—with much tighter binding than in the ordinary hydrogen atom. The resultant electrical charge of such a combination is zero, while the mass might be expected to be about that of the proton, that is, one atomic unit ($\leqq 1.7\times10^{-24}$ gm). This would give us the hint that the neutron might be considered to be a hydrogen atom in a stationary state with energy considerably below what we now consider to be the "normal" state. Langer and Rosen[4] have examined Schrödinger's wave equation to see if it would permit the existence of such a stationary state, but have concluded that so long at least as one supposes the proton and electron to act on each other with a force varying inversely as the square of the distance between them, even for very small distances, then the answer is in the negative; i.e., no such state seems possible. One may, of course, try to obtain a positive answer by making various assumptions concerning the deviation from the inverse square law of force at very small separations, but in the absence of any definite theory of the electron and the proton no very conclusive progress can be made in this direction. On the other hand, it may very well be true that we shall not arrive at a satisfactory theoretical proof the possible existence of such a neutron until we have progressed much further on the road to an unified theory of quanta and relativity. We have indeed a very vague picture of what may happen within distances of the order of 10^{-13} centimeter!

Quite apart from our troubles in showing that such a neutron *can* exist on the basis of present theories, is the question of examining the relation of the hypothesis that it *does* exist to the experimental facts, and of determining the properties which it must have in order to fit in with these experimental facts. We have seen that the picture of the neutron as a close combination of an electron and a proton would make it electrically neutral but possessing a certain gravitational mass. Since we should not wish to suppose that the electron and proton are exactly coincident in space we must admit that the system has some kind of electrical moment. In accordance with the fact that in their normal states atoms are found to be so symmetrical that they do not possess an electrical moment of first order (dipole moment) on the average, we might guess that the same would be true of the neutron, and that it would have at most a quadrupole moment. Considering that the binding between the electron and proton must be very tight, only a very small moment would be induced by an electrical field of any reasonable order of magnitude. Only a very intense electrical field would have any effect on the motion of such a neutron, but it might be that in collisions with electrons and nuclei it could act as an effective disrupting agent.[6]

It was stated above that our first guess might be that the mass of the neutron would be about the same as the mass of the proton. This is not altogether certain since we must expect that in the process of formation of the neutron considerable energy should be given off in one form or another, perhaps as radia-

[1] J. Chadwick, Nature, Feb. 27 (1932), p. 312.

[2] E. Rutherford, Proc. Roy. Soc. **A42,** 1964 (1920).

[3] W. D. Harkins, Journ. Am. Chem. Soc. **42,** 1965 (1920).

[4] R. M. Langer and N. Rosen, Phys. Rev. **37,** 1579 (1931).

[5] W. Pauli, Pasadena Meeting of the American Physical Society, June 1931; Phys. Rev. **38,** 579 (1931) (Title only).

[6] J. F. Carlson and J. R. Oppenheimer, Phys. Rev. **38,** 1787 (1931).

tion or in a triple collision with a third body, and the theory of relativity has taught us to expect that this loss of energy entails also a proportionate loss of mass. The tighter the binding of the proton and the electron in the neutron the greater will be the loss in energy in its formation and the smaller will be the final mass of the system. Just how much mass will be lost in this way cannot be predicted beforehand. To fix orders of magnitude we may say that the loss of mass in the recombination of a proton and an electron to form an atom of hydrogen in its normal state is about 2×10^{-22} gm, or about 2×10^{-5} of the mass of the electron.

In his recent letter Chadwick[1] has reported his experiments, and in a later letter to *Nature* H. C. Webster[7] of Bristol University reports very similar independent experiments with results in substantial agreement with those of Chadwick. It has been shown by previous experiments by Bothe and others that when beryllium atoms are bombarded by α-particles from polonium, a radiation of a very penetrating type is produced. If it is passed through various substances such as hydrogen, nitrogen, argon, etc. it is found that the radiation ejects particles with very high speeds which are found to be recoil atoms (or protons in the case of hydrogen). It is the character of this radiation which is in doubt. If it is assumed that the radiation consists of photons, then it is to be supposed that the recoil atoms get their energy by scattering the photons in the manner visualized in the theory of the Compton effect. From the observed energies which the recoil protons are observed to have when the radiation is passed through hydrogen, Chadwick estimates that the energy of the photons initially would be about 50×10^6 electron-volts. If this value is used for the calculation of the maximum kinetic energy which a nitrogen atom could gain by scattering a photon a value of about 4×10^5 electron-volts is obtained. Actually the recoil atoms from nitrogen are observed to have a much higher energy than this, which makes it difficult to explain the results for both hydrogen and nitrogen as a Compton scattering of photon radiation.

It is furthermore not easy to explain how the photons arising from the interaction of the α-particle and the beryllium nucleus could have an energy as high as 50×10^6 electron-volts. Assuming that the α-particle is captured by the nucleus of the Be^9 atom with the formation of a carbon nucleus:

$$Be^9+He^4\rightarrow C^{13}+\text{photon},$$

Chadwick estimates that the photon should have a maximum energy of only about 14×10^6 electron-volts.

On the other hand if one assumes that in the capture of the α-particle, a neutron is emitted and a C^{12} nucleus is formed:

$$Be^9+He^4\rightarrow C^{12}+\text{neutron},$$

then Chadwick calculates that the neutron should have a velocity of about 3×10^9 cm/sec. if it has a mass of 1 unit. A head-on collision of such a neutron with a proton would give the proton a velocity of the same order of magnitude as is observed in the recoil atoms from hydrogen.

The hypothesis of the neutron has also been put to use in another field recently. Following the suggestion of Pauli made at Pasadena last summer, a concerted effort has been made to test, both theoretically[6] and experimentally[8] the hypothesis that cosmic rays are composed of neutrons. As yet we can only say that the evidence is inconclusive.

E. Hill

[7] H. C. Webster, Nature, March 12 (1932) p. 402.

[8] L. M. Mott-Smith and G. L. Locher, Phys. Rev. **38,** 1399 (1931); L. D. Huff, Phys. Rev. **38,** 2292 (1931); W. H. Watson and F. R. Terroux, Phys. Rev. **38,** 2291 (1931).

IMPORTANT NOTICE TO AUTHORS AND SUBSCRIBERS

•

The American Institute of Physics has taken over the editorial mechanics and business management of the journal PHYSICS.

Please note carefully the addresses to which correspondence should be sent.

•

JOHN T. TATE, *Editor*
University of Minnesota
MINNEAPOLIS, MINNESOTA

will continue to receive:

Manuscripts submitted for publication

All correspondence concerning manuscripts prior to their acceptance or rejection.

•

THE EDITORIAL SECRETARY
American Institute of Physics
11 East 38th Street
New York City, New York

will receive:

All correspondence concerning the details of publication of accepted articles

All proof of articles in process of publication

All correspondence concerning reprint orders

All subscriptions and renewals

All orders for back numbers

Important Notice to Contributors

•

A JOURNAL which publishes the results of research renders a service not only to the reader but also to the author and to the institution which sponsors his work. No research in pure science may be regarded as complete until its results have been made a matter of public record. From that point of view the publication of research may be regarded as as much an integral part of the research as any other mechanical feature of it. Considerations of this kind have lead the American Physical Society, The Optical Society of America, The Acoustical Society of America and The Society of Rheology to adopt the general policy of levying a charge of $2.00 per page against the material published in their official journals of research.

In keeping with this policy a charge of $2.00 per page will be made for articles published in PHYSICS beginning with the issue of July. Bills for this amount will be sent to the authors with instructions to present them for payment to the institution supporting their work. In no case is the author expected to pay the charge himself. Should for any reason the institution be unable to make payment the bills should be returned marked "no funds available."

In partial return for this charge the journal PHYSICS will supply 100 unbound reprints of each article to the institution.